

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 24, 2004, 20:09:36 ; Search time 5876 Seconds  
(without alignments)  
4772.455 Million cell updates/sec

Title: US-10-086-464-2  
Perfect score: 3453  
Sequence: 1 MSSAPSGTGGPPSPNST.....REMEMGKIKRTGGQSGPSL 647

Scoring table: BL0SUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2/1/USPTO\_spool\_p/US10086464/runat\_23042004\_083015\_5613/app\_query.fasta\_1.839  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=epct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10086464 @CGN 1.1 5265 @runat\_23042004\_083015\_5613 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pin.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3453	100.0	1944	6	AX088876 Sequence
2	3453	100.0	1944	6	AX825703 Sequence
3	3453	100.0	2189	6	AX825705 Sequence
4	3453	100.0	2189	8	AY028699 Braesica
5	2936.5	85.0	1959	6	AX825738 Sequence
6	2936.5	85.0	1959	8	BT008400 Arabidops
7	2936.5	85.0	2098	8	AY128792 Arabidops
8	2936.5	85.0	2116	8	AY056788 Arabidops
9	2936.5	85.0	2188	8	AY059901 Arabidops
10	2936.5	85.0	2190	8	AY093065 Arabidops
11	2934.5	85.0	1959	8	BT008409 Arabidops
12	2934.5	85.0	2257	8	AF370509 Arabidops
13	2885	83.6	2324	8	AY089024 Arabidops
14	2622	75.9	79706	8	AB020746 Arabidops
15	1841.5	53.3	2307	8	AK103247 Arabidops
16	1821.5	52.8	1812	6	AX825735 Sequence
17	1812	52.5	80393	8	AP000382 Arabidops
18	1748.5	50.6	1902	6	AX088882 Sequence
19	1748.5	50.6	1902	6	AX825712 Sequence
20	1748.5	50.6	1939	6	AX088881 Sequence
21	1748.5	50.6	1939	6	AX825711 Sequence
22	1746	50.6	2013	8	BT005955 Arabidops
23	1746	50.6	2320	8	AK118488 Arabidops
24	1687	48.9	2025	6	AX088885 Sequence
25	1687	48.9	2025	6	AX825715 Sequence
26	1687	48.9	2104	6	AX088884 Sequence
27	1687	48.9	2104	6	AX825714 Sequence
28	1668	48.3	1515	6	AX825736 Sequence
29	1637.5	47.4	1488	6	AX825737 Sequence
30	1574.5	45.6	1185	8	AY224526 Oryza sat
31	1567.5	45.4	2364	8	AK064397 Oryza sat
32	1567.5	45.4	2371	8	AK102832 Oryza sat
33	1507	43.6	82356	8	AC006135 Arabidops
34	1507	43.6	92861	8	AC006439 Arabidops
35	1480.5	42.9	125021	8	AC007504 Arabidops
36	1479.5	42.8	114367	8	ATF10M10 Arabidops
37	1479.5	42.8	125502	8	ATT4L20 Arabidops
38	1479.5	42.8	195921	8	ATCHRIV81 Arabidops
39	1459.5	42.3	2158	8	AY113877 Arabidops
40	1459.5	42.3	3064	8	AY035076 Arabidops
41	1456	42.2	2046	8	AY113039 Arabidops
42	1456	42.2	2616	8	AF424623 Arabidops
43	1456	42.2	2735	8	AY075681 Arabidops
44	1434.5	41.5	120465	2	AC137001 Oryza sat
45	1431	41.4	3070	8	AK110196 Oryza sat

ALIGNMENTS

RESULT 1

AX088876  
 LOCUS AX088876 1944 bp DNA linear PAT 17-MAR-2001  
 DEFINITION Sequence 1 from Patent WO0114563.  
 ACCESSION AX088876  
 VERSION AX088876.1 GI:13397639

KEYWORDS  
 SOURCE Brassica napus (rape)

ORGANISM

Brassica napus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE

AUTHORS Goring, D. and Silva, N.

TITLE Proline-rich extensin-like receptor kinases

JOURNAL Patent: WO 0114563-A 1 01-MAR-2001;

Goring, Daphne (CA) ; Silva, Nancy (CA)

FEATURES

Location/Qualifiers

source

1..1944

/organism="Brassica napus"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3708"

CDS

1..1944

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAC34703.1"

/db\_xref="GI:13397640"

/translation="MSSAPSGTGSPPSPNSNTTTPPASAPPTTPSPPPPTI

PTSPPPSSRTSPASPPSPSPPLPQSPAPPTPGSPPPAPVPTPNPP  
 VSPGPPNSRSGSPSPSPSPSSDGLSGVVGIAIGVALLVITLCLC  
 KKKRRDEADYYVPPPPGPKAGFYGGQQWROQNAITPPSDVHTVSLPPPKAP  
 SPQPQPPPPPPMSSSGSDYSDRVLPPSPGLVGFSGSKSTFTYEELARATNGFS  
 EANLLGGGGYVHKGLVPSGKVAVKLVGQGGREFQAEVEIIISRVHRLHVL  
 VGYCIAGKRLVVEFVNNLELHLHGEGRPTMEWSTRLKIALGSAKGLSYLHEDCN  
 PKIITRDIKASNLIIDFKFAKVAADFGLAKIADTNTHTVSVMTGTFGLAPYAAGG  
 KLTEKSDVFGVLLLELTGRPPVDANNVVDLSLDVMDARPLLNRAEQDPEGLAD  
 AOMNNGIDREMAKVAACAACVHRHARRRPMISOIVRALEGNVSLDLNEGMRPGQS  
 NVYSYGGSDTDYSSQNMKFRKRMALGTQEYNATGEVSNPTSDYLGYSGSSSEG  
 QTTREMEMGKIKRTGGYSGPSL"

ORIGIN

Alignment Scores:

Pred. No.: 4,02e-76 Length: 1944

Score: 3453.00 Matches: 647

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-086-464-2 (1-647) x AX088876 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20  
 Db 1 ATGTCCTCGGCGCGTCTCCGGGAC TGGTTCGCTCCATCTCCACCATCAAACTCCACA 60  
 Qy 21 ThrThrThrProProAlaSerAlaProProThrThrProSerSerProProPro 40  
 Db 61 ACCACCACTCTCTCCAGCTTCGGCTCTCTCTCCACCACTCTCTCTCTCTCTCTCCGCG 120  
 Qy 41 ProSerThrileProThrProProProProSerSerSerArgSerThrProSerAlaProPro 60  
 Db 121 CCATCCCACTATCCGACATCT 180  
 Qy 61 ProSerProProThrProSerThrProGlySerProProProLeuProGlnProSerPro 80  
 Db 181 CCATCTCCACCACTCCATCTACGCGGGAATCTCCACCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
 Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100  
 Db 241 CCCGCTCCAACTACGCGCGGATCTCCACCCGACCTGTGTACTCTCTCTCTCTCTCTCTCTCT 300  
 Qy 101 ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro 120

Db 301 CCACCTTCAGTCCAGGACACCGTCCAAATCCTTTCACGCGAAGGAGGATCTCTCTCGACCT 360  
 Qy 121 ProSerSerProSerProProSerProSerSerSerAspGlyLeuSerThrGlyValValVal 140  
 Db 361 CCATCTTCTCCCTCGCGCGCTCTCTCTTCTTCCAGCGGTTATCAACAGGAGTGTGTGTG 420  
 Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCyeLeuLeuCys 160  
 Db 421 GGAATCCCCATCGGAGGAGTCTCTCTCTGTGTAGTAGTCTCTGATTTGTCTCTCTCTGT 480  
 Qy 161 LysLysLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProProPro 180  
 Db 481 AAGNAGAAACGACGAGAGAGAGAGAGAGTCTTACTATGTCTCTCGCCACCTCTCTCT 540  
 Qy 181 GlyProLysAlaGlyGlyProProGlyGlyGlnGlnGlnGlnGlnTrpArgGlnGlnAsnAla 200  
 Db 541 GGTCCCAAGCCGAGGACCTTACGGTGGACAGCAACATGGCGGCAACAACAAACGCA 600  
 Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProLysAlaProSerPro 220  
 Db 601 ACACACCGTCAGATCATGTGTCGTGACGTCTACTACCAACACCATAGGCTTCAATCTCCA 660  
 Qy 221 ProArgGlnProProProProProProPheMetSerSerSerGlyGlySerAsp 240  
 Db 661 CCACGGCAACCTCTCTCCACCTCCACACCGCTTTCATGAGCAGCAGCGCGCTCCGAC 720  
 Qy 241 TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys 260  
 Db 721 TACTCGGACCGTCAGTTCTCTCCACCGCTCTCCAGGCTTGTGTAGGCTTCTCCAAA 780  
 Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280  
 Db 781 AGCACTTTCATACGAGGAGCTAGCTAGAGCCACCAATGGTTTCTCCGAGCGCAACTTG 840  
 Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300  
 Db 841 TTAGGACACGCGGTTTGGTTACGTGCACAAAGGTGTCTTGTGTGTATTTGATGATGCC 900  
 Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320  
 Db 901 GCTGTGAAGCAGTTGAAAGTTGGAGTGGTTCAGGAGAGAGGAGTTTCAGGACAGAGTT 960  
 Qy 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340  
 Db 961 GAGATCATCAGCAGAGTTCCACACAGGCGATCTGGTGCTCTTGTGTGTATTTGATGATGCC 1020  
 Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360  
 Db 1021 GGTGCCAAAGATTTGCTTGTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACTC 1080  
 Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380  
 Db 1081 CATGGCAGGAGCGCGCTTCAATGGAAATGGAGCACCAGATTGAAGATTGCTCTTTGGATCT 1140  
 Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400  
 Db 1141 GCTAAAGGACTTCTTATCTTCATGAAGATTGCAATCTTAATCATTCACCGGTATATC 1200  
 Qy 401 LysAlaSerAsnIleLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420  
 Db 1201 AAGCTTCAACATATTGATAGATTTCAGTTTGAAGCTAAGGTTGCTGATTTTGTCTT 1260  
 Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440  
 Db 1261 GCTAAGATTGCTCTGATACAAACACGCATGTATCAACAGTGTGTGATGGGAACCTTTGGG 1320  
 Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460  
 Db 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCAGGAGAGTCTGAGCTTTTCTCA 1380  
 Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnVal 480  
 Db 1381 TTTGGCGTGTGCTTTTGGAGCTCATTTACTTGGACGTGACCGCTGTGATGCCCAACATGTC 1440



Db 841 TTAGCAAGCGGGTTCGGTTACGTGCACAAAGGTGTGTTCCCTAGTGGGAAGAAGTT 900  
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320  
Db 901 GCTGTGAAGACAGTTTGAAGATTGGGAGTGGTTCAGGAGAGAGAGAGATTTCAGGACGAGGTT 960  
Qy 321 GluIleIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340  
Db 961 GAGATCATCAGCAGAGTTTACACACAGGCATCTGGGTCTCTGTGTGGTTATTCGATCC 1020  
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360  
Db 1021 GGTGCCAAAGATTGCTTGTCTATGAGTTTGTCTTCAACAATCTCGAGCTTCACCTC 1080  
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380  
Db 1081 CATGCGGAGGAGCGCCACATGAAATGGAGCACAGATTCAAGATTGCTCTTGGATCT 1140  
Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400  
Db 1141 GCTAAAGGACTTCTTATCTTCATGAAGATTGCAATCTTAAATCATTCACCGGTATATC 1200  
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420  
Db 1201 AAGGCTTCACACATATTGATAGATTCAAGTTTGAAGCTAAGTTGCTGATTTTGGTCTT 1260  
Qy 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440  
Db 1261 GCTAAGATTGCTTCTGATACAAACACGATGTATCAACACGCTGTGATGGGAACCTTTGGG 1320  
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460  
Db 1321 TACTTGGCTCCGGAATAGCTGCAAGCGGAAGCTCACGGAGAAGTCTGAGCTTTTCTCA 1380  
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAlaAsnVal 480  
Db 1381 TTTGGCGTTGTCTTTGGAGCTCATTACTGGAGCTCGACCGTGTGATGCCAACAATGTC 1440  
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500  
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGCGACGACCATTTGCTTAAACCGAGCATCTGAGCAA 1500  
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMet 520  
Db 1501 GGAGACTTTGAGGGTTTACCTGATGCCAAGATGAATAATGGGTATGACAGAGAGAGATG 1560  
Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540  
Db 1561 GCTCGCATGGTTGCTTGTGCTGGGCTTGTGTTCGCCATTTCAGCTCGCGCGAGACCTGCG 1620  
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560  
Db 1621 ATGAGCCAGATTGTGCGTGGCTTAGAAGAAATGTATCACTGTCCAGATCTTTAAACGAAGGG 1680  
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580  
Db 1681 ATGAGCCAGATTGTGCGTGGCTTAGAAGAAATGTATCACTGTCCAGATCTTTAAACGAAGGG 1740  
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600  
Db 1741 AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAAATGGCATTGGAACCTCAAGAGTAC 1800  
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620  
Db 1801 AACGCCAGGGTGAGTACAGTAATCCGACCATGTACTATGGACTGTATCCCGTGTGTTCA 1860  
Qy 621 SerSerGluGlyGlnThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640  
Db 1861 AGCAGGAGGGCCAAACACACCGGAAATGGAGATGGGGAAGATTAAAGAACCGGTGAG 1920  
Qy 641 GlyTyrSerGlyProSerLeu 647  
Db 1921 GGTATATGAGCCCTTCTCTT 1941

## RESULT 3

AX825705 LOCUS AX825705 2189 bp DNA linear PAT 11-DEC-2003  
DEFINITION Sequence 3 from Patent WO03072763.  
ACCESSION AX825705

VERSION AX825705.1 GI:39751232

KEYWORDS Brassica napus (rape)

SOURCE Brassica napus

ORGANISM Brassica napus

REFERENCE 1  
AUTHORS Goring, D., Silva, N. and Haffani, Y. Z.  
TITLE Increasing plant seed production  
JOURNAL Patent: WO 03072763-A 3 04-SEP-2003;  
Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)

FEATURES  
source 1..2189  
/organism="Brassica napus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3708"

ORIGIN

Alignment Scores: 4.39e-76 Length: 2189

Pred. No.: 3453.00 Matches: 647

Score: 3453.00 Conservative: 0

Percent Similarity: 100.00% Mismatches: 0

Best Local Similarity: 100.00% Indels: 0

Query Match: 100.00% Gaps: 0

DB: 6

US-10-086-464-2 (1-647) x AX825705 (1-2189)

Qy 1 MetSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20

Db 97 ATGTCTCGGCGCGCTCTCGGGAGCTGGTTCGCTCCATCTCCACCATCAATCCACA 156

Qy 21 ThrThrThrProProAlaSerAlaProProProThrProSerSerProProPro 40

Db 157 ACCACCACTCTCTCCAGCTTCGCTCTCTCCACCAACCTCTCTCTCTCTCTCT 216

Qy 41 ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaPro 60

Db 217 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276

Qy 61 ProSerProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80

Db 277 CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCT 336

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100

Db 337 CCCGCTCAACTACGCGCGGATCTCCACCGCGACCTGTACTCTCTCTCTCTCTCT 396

Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120

Db 397 CCACCTTCAGTCCAGGACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 456

Qy 121 ProSerSerProProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140

Db 457 CCATCT 516

Qy 141 GlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160

Db 517 GGAATCGCATCGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576

Qy 161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProPro 180

Db 577 AAGAAGAAACACGAGAGAGACGAAGAAGATGCTTACTATGTCTCTCTCTCTCTCTCT 636

Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 200



DB: 8 Gaps: 0  
US-10-086-464-2 (1-647) x AY028699 (1-2189)  
Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20  
Db 97 ATGTCCTCGGCGCGCTCTCCGGGACTGGTTCGCCTCCATCTCCACCATCAAACTCCACA 156  
Qy 21 ThrThrThrProProProAlaSerAlaProProProThrThrProProSerSerProPro 40  
Db 157 ACCACCACTCTCTCTCCAGCTCCGCTCTCTCTCCACACACACCTCTCTCTCTCCGCGC 216  
Qy 41 ProSerThrIleProThrSerProProSerProSerSerArgSerThrProSerAlaProPro 60  
Db 217 CCATCCACTATTTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276  
Qy 61 ProSerProProThrProSerThrProGlySerProProProProProLeuProGlnProSerPro 80  
Db 277 CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCCA 336  
Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100  
Db 337 CCGCTCCCAACTACGCGCGGATCTCCACCGCACCTGTACTCTCTCTCTCTCTCTCTCTCT 396  
Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120  
Db 397 CCACCTTCAGTCCCAAGSACCAACCGTCCATCTCTCCACGGAAGGAGATCTCTCTCGACCT 456  
Qy 121 ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140  
Db 457 CCATCTCTCTCTCGCGCGCT 516  
Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160  
Db 517 GGAATCGCCATCGGAGGAGTCTCTGCTGTGATAGTACTCTGATTTGTCTCTCTCTCT 576  
Qy 161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProProPro 180  
Db 577 AAGAAGAAACACGAGAGAGAGAAAGATGCTTACTACTATGTTCCTCCGCCACCTCTCTCT 636  
Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnTrrArgGlnGlnAlaAsnAla 200  
Db 637 GGTCCCAAAACCGGAGGACCTTACGTTGGA CAGCAGCAACATGGCGGCAACAAACGCA 696  
Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProAlaProSerPro 220  
Db 697 ACACCACCGTCAGATCATGTGTGACGTCACTTACCACCAACCACTTAAGGCTCCATCTCCA 756  
Qy 221 ProArgGlnProProProProProProProProPheMetSerSerSerGlyGlySerAsp 240  
Db 757 CCACGGCAACCTCTCCACCTCCACCGCTTTCATGAGCAGCAGCGCGGCTCCGAC 816  
Qy 241 TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys 260  
Db 817 TACTCGGACCGTCCAGTCTTCTCTCCACCGCTCCAGGGCTGTGTAGGCTTCTCCAAA 876  
Qy 261 SerThrPheThrTyrGluGluLeuAlaAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280  
Db 877 AGCACTTTCCATACAGGAGCTAGCTAGAGCCCAATGGTTTCTCCGAGGCGAATCTG 936  
Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProProSerGlyLysGluVal 300  
Db 937 TTAGGACAAGCGGGTTCGGTTACGTGCACAAAGGTGTGTTCCTAGTGGGAAAGAAAGTT 996  
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320  
Db 997 GCTGTGACAGCAGTTGAAGATTGGGAGTGTTCAGGGAGAGAGGAGTTTCAGCGCAGAGTT 1056  
Qy 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340  
Db 1057 GAGATCATCAGCAGAGGTTCCACACAGGCATCTGGTGTCTCTGTGTTTATTTGATCGCC 1116  
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuLeuHisLeu 360

Db 1117 GGTGCCAAAGATTCCTTGCTATGAGTTGTTCCTAACCAACATCTCGAGCTTCACCTC 1176  
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380  
Db 1177 CATGGCAGGAGCGCCTACAATGGAATGGAGCACCAAGATTGAAGATTGCTCTTTGGATCT 1236  
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400  
Db 1237 GCTAAAGAGCTTCTTATCTTCATGAAGATTGCAATCCTAAAATCATTTACCGTGATATC 1296  
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420  
Db 1297 AAGGCTTCAACATATTGATAGATTTCAGTTTGAAGCTAAGGTTCGTGATTTGGTCTT 1356  
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440  
Db 1357 GCTAAGATTGCTTCTGATACAAACACGATGATCAACACGCTGTGATGGGAACCTTTGGG 1416  
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460  
Db 1417 TACTTGGCTCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTCA 1476  
Qy 461 PheGlyValValLeuLeuLeuIleThrGlyArgArgProValAspAlaAsnVal 480  
Db 1477 TTTGGCGTTGCTTTTGGAGCTCATTTACTGGAGCTGACCCGTTGATGCCCAACATGTC 1536  
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500  
Db 1537 TATGTAGATGACAGCTTAGTTGACTGGGCACGACCAITTCCTTAACCGAGCATCTGAGCAA 1596  
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520  
Db 1597 GGACACTTTGAGGGTTAGCTGATGCAAGATGAATAATGGGTATGACAGAGAGAGATG 1656  
Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540  
Db 1657 GCTCGCATGGTTGCTGTGCTGCGGCTGTGTTGCCATTCAGCTCGCCGCGAGACCTCGC 1716  
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560  
Db 1717 ATGAGCCAGATTGTGCTGCGTTAGAAGAAATGTATCATCTGTGATGATTAACGAAGGG 1776  
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580  
Db 1777 ATGAGACAGGTCAAGCAATGTATACAGCTCATACGGAGGAAGACCGATTATGACTCG 1836  
Qy 581 SerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600  
Db 1837 AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAATGGCACTTGGAACTCAAGAGTAC 1896  
Qy 601 AsnAlaThrGlyGlyTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620  
Db 1897 AACGCCACGGGTAGTACAGTAATCCGACCACTGACTATGAGCTGTATCCCGTCTGGTTCA 1956  
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640  
Db 1957 AGCAGCAGGCGCCAAACACACACCGGAAATGGAGATGGGGAAGATTAAAGAAACCGGTG 2016  
Qy 641 GlyTyrSerGlyProSerLeu 647  
Db 2017 GGTTATAGTGACCTTCTCTT 2037  
RESULT 5  
AX825738  
LOCUS AX825738 1959 bp DNA linear PAT 11-DEC-2003  
DEFINITION Sequence 36 from Patent WO03072763.  
ACCESSION AX825738  
VERSION AX825738.1 GI:39751255  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

1 Goring, D., Silva, N. and Haffani, Y. Z.

TITLE Increasing plant seed production

JOURNAL Patent: WO 03072763-A 36 04-SEP-2003;

Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, 2. (CA)

Location/Qualifiers

FEATURES

source

1. 1959

/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3702"

## ORIGIN

Alignment Scores:

Pred. No.: 1-23e-63 Length: 1959

Score: 2936.50 Matches: 574

Percent Similarity: 90.61% Conservative: 24

Best Local Similarity: 86.97% Mismatches: 41

Query Match: 85.04% Indels: 21

DB: 6 Gaps: 12

US-10-086-464-2 (1-647) x AX825738 (1-1959)

Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17  
Db 1 ATGTCCACAGCGCGTCTCCAGGCACCTACTCCATCACCCTCCACCGTCTCTCCACCA 60  
Qy 18 AnSerThrThrThrThrProProProLa-----SerAlaProProThrThrPro 35  
Db 61 AACTCGCAACACCACTCTCTCCAGCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
Qy 36 SerSerProProProProSerThrThrThrThrThrThrThrThrThrThrThrThr 54  
Db 121 TCTTCT 180  
Qy 55 ThrProSerAlaProProProSerProProThrProProThrProProGlySerProPro 74  
Db 181 CCT 228  
Qy 75 LeuProGlnProSerProProAlaProThrThrThrProGlySerProProAlaProVal 94  
Db 229 CTTCCTCAACCTTCCCGCTCGCTCCATCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 285  
Qy 95 ProProThr---ArgAnProProProSerValProGlyProProProSerAnProSer 113  
Db 286 CCCTCAAAACCTCGAAGCCCTCCATCTCTTAACCAAGGACCCACCAACATCTCCCTCA 342  
Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSer 131  
Db 343 ---GGATCTACTCTAGAACCTCCATCAACACACTAAACCGTCCGCGCTCTGATTCTCC 399  
Qy 132 AspGlyLeuSerThrGlyValValValGlyLeuAlaLeuGlyValAlaLeuVal 151  
Db 400 GATGGATTGTCTACCGGAGTTGTGTAGGATCGCCATTTGGAGGAGTCTGATTCTTGT 459  
Qy 152 IleValThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169  
Db 460 ATACTGACTCTGATTGTCT 519  
Qy 170 AspAlaTyrTyrValProProProProProProProProProProProProProPro 189  
Db 520 GCTGTTACTATGTTCT 579  
Qy 190 GlyGlnGlnGlnGlnThrArgGlnGlnAnAlaThrProProSerAsp---HisValVal 208  
Db 580 GGTCAACAACAGTATTGGCAACAAACACCGCTCGCGCGCTCAGATAATCATGTAGTG 639  
Qy 209 ThrSerLeuProProProLeuAlaProSerProProProProProProProProProPro 228  
Db 640 ACGTCATTG---CCACCACCTTAAGCCTCCATCTCCACCAACCAACCTCTCTCCGCC 696  
Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247

Db 697 CCACCACAGCAATTCATAGTAGCAGTGGTGTCTGACTATTCCGATCTTCCGGTCTT 756  
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267  
Db 757 CTCCACCATCTCCAGGCTTGTGTAGGCTTTCTAAAGACATTTTACATTTATGAGAG 816  
Qy 268 LeuAlaArgAlaThrAnGlyPheSerGluAlaAnLeuLeuGlyGlnGlyGlyPheGly 287  
Db 817 TTGTCGAGAGCTACTAATGCTTCTCTGAGGCTAATTTGTAGGCAAGGAGGCTTGGT 876  
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuVal 307  
Db 877 TATGTCATAAAGGTATATTCCTAGTGGGAAAGTCTGCTGTAACAGATTTCAAAGCT 936  
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuLeuValHis 327  
Db 937 GGTAGTGGTTCAGGAGAGAGAGAGTTCAGGCTGAGGTGTGAGATCATTAGCAGAGT 996  
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347  
Db 997 CACAGGCAATTTGGTTCCTTATTTGGTATTGTTATGTCGGCGGTGTTCAAAGATTT 1056  
Qy 348 TyrGluPheValProAnAnAnLeuLeuLeuHisGlyGlyGlyGlyGlyGlyGlyGly 367  
Db 1057 TATGAGTTTGTTCACAAACAATCTTGTAGTTCACCTCCATCGTAAGGCGTCAACAT 1116  
Qy 368 MetGluTrpSerThrArgLeuValLeuAlaGlySerAlaLysGlyLeuSerTyrLeu 387  
Db 1117 ATGGAATGGAGTACTAGATTGAAGATTGCTCTGGATCTCTTAAGGACATTTTCAAT 1176  
Qy 388 HisGluAspCysAnProLysIleLeuHisArgAspIleLysAlaSerAnIleLeu 407  
Db 1177 CATGAAGATTGCAATCCGAAATCATTCACCGTGATTAAGGCGTCAACATATTGAT 1236  
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 427  
Db 1237 GATTTCAAATTTGAAGTAAAGTTGCTGCTTTGGTCTTCCCAAGATTTGCTCTGATA 1296  
Qy 428 AnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447  
Db 1297 AACATCATGATATCTACACGCTGATGGGAACCTTTGGGTATTTGGCTCCGGAATA 1356  
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeu 467  
Db 1357 GCAAGTGAAGCTCACAGAAAAGTCTGAGCTTTCTCATTTGGCGTGTACTTTTGG 1416  
Qy 468 LeuIleThrGlyArgArgProValAspAlaAnAnValTyrValAspAspSerLeuVal 487  
Db 1417 CTTATTACTGGAGGCGCCCTGTTGATCGAACAAATGCTATGTATAGATGACAGCTTAG 1476  
Qy 488 AspTrpAlaArgProLeuLeuAnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507  
Db 1477 GACTGGGCACGACCATTTGCTTAACCGGACATCTGAGGAGGAGATTTTGAAGGTT 1536  
Qy 508 AspAlaLysMetAnAnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527  
Db 1537 GATTCAAAGATGGTAAATGAGTATGACAGAGAGAGATGGCTCGCATGGTGTGGCT 1596  
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgProArgMetSerGlnIleValArg 547  
Db 1597 GCGGCTGTGTCGCCATTGCTGCTCGCGCAGACCTCGCATGAGCCAGATAGTACGGCG 1656  
Qy 548 LeuGluGlyAnValSerLeuSerAspLeuAnGluGlyMetArgProGlyGlnSerAn 567  
Db 1657 TTAGAAGGAAATGTATGCTGTCTGATCTTAAAGGAGGATGAGACCGGGTCAAGCAAC 1716  
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAnGluAspMet 587  
Db 1717 GTATACAGCTCATATGGAGGAGCAGACAGACTATGACACGAGCCCATACACGAGCAT 1776  
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAnAlaThrGlyGlyTyrSer 607

Db 1777 ATAAAGTTTAGGAAATGGCTCTTGGAACTCAAGATAAGCGCACAAACCGCGAGTACAGT 1836  
 Qy 608 AenProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGlyGlnThrThr 627  
 Db 1837 AATCCAACCACTGACTACGAGCTGTACCGCTCTGGTTCAAGCAGTGAAGGTCAAGCCACA 1896  
 Qy 628 ArgGluMetGluMetGlyValIleValArgThrGlyGlnGlyTyrSerGlyProSerIleu 647  
 Db 1897 CGAGAAATGGAGATGGGAAGATTAAAGAAACCCGGTCAAGGTTATAGTGGACCCCTCTCTT 1956

## RESULT 6

BT008400 1959 bp mRNA linear PLN 15-MAY-2003  
 LOCUS Arabidopsis thaliana A3924550 gene, complete cds.

DEFINITION BT008400

ACCESSION BT008400.1 GI:30725473

VERSION FLI CDNA.

KEYWORDS Arabidopsis thaliana (thale cress)

## SOURCE

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 1959)

Chen, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P.,

Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T.,

Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L.,

Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C.,

Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,

Theologis, A. and Ecker, J.R.

Arabidopsis ORF clones

Unpublished

2 (bases 1 to 1959)

Chen, R., Chen, H., Kim, C.J., Shinn, P., Bowser, L., Carninci, P.,

Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T.,

Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M.,

Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L.,

Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C.,

Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W.,

Theologis, A. and Ecker, J.R.

Direct Submission

Submitted (15-MAY-2003) Salk Institute Genomic Analysis Laboratory

(STGenAL), Plant Biology Laboratory, The Salk Institute for

Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,

USA

The discrepancy does not affect the protein sequence.

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN

Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,

Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,

Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and

sequenced the pUNI (ORF) clones using the RAFL cDNAs: Chen, R.,

Chen, H., Kim, C.J., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M.,

Dale, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,

Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,

Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,

Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and

Ecker, J.R.

Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)

contributed equally to this work as PIs.

## FEATURES

source

1. 1959

Location/Qualifiers

/organism="Arabidopsis thaliana"

/molecule="mRNA"

/db\_xref="taxon:3702"

/chromosome="3"

/clone="U17522"

/ecotype="Columbia"

/note="This clone is in pUNI 51"

## CDS

1. 1959  
 /note="unknown protein"  
 /codon\_start=1  
 /product="At3g24550"  
 /protein\_id="AAP37759.1"  
 /db\_xref="GI:30725474"  
 /translation="MSTAPSGTTPSPSPPTNSTTTTPPPAASPPPTTTSSPP  
 PSSTNSTSPSPSPPLPSPSPSGSLTPPLPQSPSAPLTPSPSPPTSPSPSP  
 PSNQGPNTSPSGTTPSTPSTPSPSPSDSLSTGVVGVGIALGGVAILVILILCL  
 LCKRRRRHDEAAAYVPPPPSGGKAGPYGGQQYWOQGNASRPSDNHVTSLPP  
 PKPPSPRPKPPPPPPAFMSSSGSDYSLPLVPPSPGLVLFSGSKSTFTYELSLRA  
 TNGVSEANLLGQGGFYVHKGIILPSGKEVAVKOLKAGSGQGERFOAEVEIISRVHHR  
 HLVSILGYCWAGVQRLVAVFVNNLEFLHKGKRPMTWSTRLKIALGSAKGLSYL  
 HEDCNPKLIHRDIKASNILIDFPEAKVADFGLAKIASDNTNHTVSTVMGTFFGYLAPE  
 YAAAGKUTEKSDVFSFGVLELLITGRPPVDANNVYVDDSLVDWARFLNRSBEGDF  
 EGLASXKGNEYDREEMARMVACAACVCAARRRPRMSQLVRALEGNVLSLDNEG  
 RPHGSHVSYSGYSTDYDTQYNDMDIKFRKMAFGTQEGYGTGTSNPTSDGYLPSG  
 SSSEGGQATREMENGKIKKTGGYSGPSL"  
 misc\_difference 1920  
 /note="compared to genome"  
 /replace="t"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.23e-63 Length: 1959  
 Score: 2936.50 Matches: 574  
 Percent Similarity: 90.61% Conservative: 24  
 Best Local Similarity: 86.97% Mismatches: 41  
 Query Match: 85.04% Indels: 21  
 DB: 8 Gaps: 12

US-10-086-464-2 (1-647) x BT008400 (1-1959)

Qy 1 MetSerAlaProSerProGlyThr-----GlySerProProSerProSer 17  
 Db 1 ATGTCCACAGCGCGTCTCCAGGCACTACTCCATCCATCCACCGTCTCTCCACCA 60  
 Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrPro 35  
 Db 61 AACTCGAACACCACCTCTCTCCAGCAGCCTCTCTCTCTCCACCAACCTCT 120  
 Qy 36 SerSerProProPro-----ProSerThrThrThrSerProProSerSerArgSer 54  
 Db 121 TCTTCTCTCTCGTCCGCGCTCACTAATCAACCTCTCTCTCTCTCTCTCTCTTTA 180  
 Qy 55 ThrProSerAlaProProProProProProProProProProGlySerProProPro 74  
 Db 181 CCT 228  
 Qy 75 LeuProGlnProSerProProAlaProThrThrThrProGlySerProProAlaProValThr 94  
 Db 229 CTTCTCAACCTTCCCTCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 285  
 Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113  
 Db 286 CCCTCAACCTCGAAGCCCTCCATCTCTCAACCAAGGACCACCAACACCTCTCTCTCA 342  
 Qy 114 GluGlySerProArgProProProSerSer-----ProSerProProSerProSerSer 131  
 Db 343 ---GGATCTACTCTAGAACTCCATCAAACTAAACCGTCGCGCGCTCTGATCTTCTCC 399  
 Qy 132 AspGlyLeuSerThrGlyValValGlyIleAlaIleGlyGlyValAlaLeuVal 151  
 Db 400 GATGGATGTTCTACCGAGTTGTGTAGGAATCCCATTTGGAGGAGTCGCTATTCTTGT 459  
 Qy 152 IleValThrLeuIleCysLeuLeuCysIleCysIleCysIleCysIleCysIleCysIle 169  
 Db 460 ATACTGACTCGATTGTTCTTCTCTGTAAAGAAACGAAGAACAGACACGACGATGAA 519  
 Qy 170 AspAlaTyrTyrValProProProProProProProProProProGlyProLysAlaGlyGlyProTyrGly 189  
 Db 520 GCTGCTTACTATGTTCT 579

Qy 190 GlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208  
 Db 580 GGTCAACAAGTAGTATTGGCAACAACAAACCGGTACCGCGCTCAGATAATCATGTAGTG 639  
 Qy 209 ThrSerLeuProProProProProProProProProProProProProProProProPro 228  
 Db 640 ACGTCATTG---CACACCCTTAAGCCTCCATCTCCACCAGAAAACCTCTCCGCCACCT 696  
 Qy 229 ProProPro---PheMetSerSerGlyGlySerAspTyrSerAspArgProValLeu 247  
 Db 697 CCACCACCATTCATCATGATGAGTGGTGGTCTGACATATTGGATCTTCGGTCTTT 756  
 Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267  
 Db 757 CCTCCACCATCTCCAGGGCTTGTGTAGCTTTTCTAAAGACACTTTCACATATGAGGAG 816  
 Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287  
 Db 817 TTGTGAGAGACTACTAATATGGCTTCTCTGAGGCTAATTTGTAGGACCAAGGAGGTTGGT 876  
 Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307  
 Db 877 TATGTGCATAAAGGTATATTCCTAGTGGGAAAGAGTGTCTGTGAACAGTTCGAAGCT 936  
 Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327  
 Db 937 GGTAGTGGTCAGGAGAGAGAGAGTTCAGGCTCAGGTTGAGATCATTTAGCAGAGTTTCAT 996  
 Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuVal 347  
 Db 997 CACAGGCATTTGGTCTCTTATTGGTATTGTATGGCCGTGTTCAAAGATTAATTCTGTC 1056  
 Qy 348 TyrGluPheValProAsnAsnLeuLeuGluLeuHisGlyGlyGluGlyArgProThr 367  
 Db 1057 TATGAGTTGTTCACAAACAACAATCTTGAGTTTACCTCCATGTGTAAGGACGCCCTACG 1116  
 Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387  
 Db 1117 ATGGAATGGAGTACTAGATTCAAGATTGCTCTTGATCTGTCTAAAGGACTTTCATATCTT 1176  
 Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407  
 Db 1177 CATGAAGATTGCATTCGAAATCATTCACCGCTGATATTAAAGGCGTCAAAACATATTGAT 1236  
 Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427  
 Db 1237 GATTTCAATTTGAAGCTAAGGTTGCTGACTTTGGTCTTGCCAGATGCTCTCTGATACA 1296  
 Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447  
 Db 1297 AACACTCATGTATCTACACGGCTGATGGGAACCTTTGGGTATTTGGCTCCGGAATATGCT 1356  
 Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467  
 Db 1357 GCAAGTGGAAAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTGGAA 1416  
 Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspSerLeuVal 487  
 Db 1417 CTTATTACTGGAGGCGCCCTGTTGATCGGAACATATGCTATGATGATGACAGCTTAGTT 1476  
 Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507  
 Db 1477 GACTGGGCACGACCATTCCTTAACCGACATCTCAGGAAGGAGATTTTGAGGTTTGCT 1536  
 Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527  
 Db 1537 GATTCAAGATGGGTAAATGAGTATCACAGAGAGAGATGGCTCGCATGGTTCCTTGCGCT 1596  
 Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547  
 Db 1597 GCGGCTGTGTCCCATTCAGCTCGCCGACAGCTCGCATGACGACAGATAGTACGGGCG 1656  
 Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567

Db 1657 TTAGAAGCAATGATCGCTCTGATCTTAACCAAGGAGTACAGCCGGTTCACAGCAAC 1716  
 Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587  
 Db 1717 GTATACAGCTCATATGAGGAAGCACAGACTATACACGAGCCCAATCAACGACGACATG 1776  
 Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSer 607  
 Db 1777 ATAAAGTTTAGGAAATGGCTCTTGGAACTCAAGATACGCACACCGCGGAGTACAGT 1836  
 Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627  
 Db 1837 AATCCAAACCATGCTACTACGAGCTGTACCCGCTCTGGTTCAAGCAGTGAAGTCAAGCCACA 1896  
 Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647  
 Db 1897 CGAGAAATGAGATGGGAAAGATCAAGAAACCGGTCAAGGTTATAGTGGACCTCTCTT 1956

RESULT 7  
 AV128792  
 LOCUS  
 DEFINITION Arabidopsis thaliana protein kinase-like protein mRNA, complete cds.  
 ACCESSION AV128792  
 VERSION AV128792.1 GI:22136227  
 KEYWORDS FLI CDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 2098)  
 AUTHORS Tripp,M., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Hayashizaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-JUL-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA  
 COMMENT e-mail for correspondence: arab@sequence.stanford.edu  
 The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.  
 The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Tripp,M., Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Tripp,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.  
 FEATURES  
 source  
 1..2098  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="U16568"  
 /ecotype="Columbia"  
 /note="This clone is in pENTR/SD-dtopo This is a cloned PCR product using RIKEN clone RAFL09-06-A17 (AY059901) as a template"  
 1..1959  
 /codon\_start=1

CDS







CDS

```

32. .1990
/ gene="MOB24.13"
/ codon_start=1
/ product="protein kinase-like protein"
/ protein_id="AAL24383.1"
/ db_xref="GI:16649063"
/ translation="MSTAPSGTTPSPSPPTNTSTTTTPPAASSPPPTTTPSSPP
PSPSTNTPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
PSNQGPPNTSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP
LCKKRRRRHDEAYVPPPPSPGKAGGPGYGGQQYQWQOONASRPSDNHVTSLPP
PKPSPRPKPPPPPPAFMSSSGSDYSLDLPVPPSPGLVGFSTFYIELSRA
HNGSEANLGGVGLYVHKGILPSGKEVAVKQKAGSGQGERFQAEVEIISRVHR
HLVSLIGCMAGQGLLYIEFVPPNNLEFHLHGKRPMTWSTRLKIALGSAKLSL
HEDCNPKIHRDIKASNLIDFKFAKVADEFLAKIASDITNTHVSTVRMGTFGLAPE
YAAAGKLTKEKSDVFSFVGLLELITGRPPVDANNVYDDSLVDWARPLLNRASEGF
EGLADSKMGNVYDREEMARMVACAAACVHSARBRPMQIIVRALEGNVSLSDLNEMG
RPSHSNVYSSVGGSTDDTDSQYNDMDMIFKPKMALQTEYGTGTEYSNPTSDYGLYPSG
SSSEGGQATREMEMGKIKKTGGYSGPSL"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1,33e-63 Length: 2188
Score: 2936.50 Matches: 574
Percent Similarity: 90.61% Conservative: 41
Best Local Similarity: 86.97% Mismatches: 44
Query Match: 85.04% Indels: 21
DB: 8 Gaps: 12

```

US-10-086-464-2 (1-647) x AV059901 (1-2188)

```

Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
Db 32 ATGTCACAGCGCGTCTCCAGGCACTACTCCATCACCATCTCCACCGTCTCTCCACCA 91
Qy 18 AsnSerThrThrThrProProProAla-----SerAlaProProThrThrPro 35
Db 92 AACTCGAACACCACTCTCTCCAGCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 151
Qy 36 SerSerProProPro---ProSerThrileProThrSerProProSerSerArgSer 54
Db 152 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 211
Qy 55 ThrProSerAlaProProSerProProThrProThrProSerThrProGlySerProPro 74
Db 212 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 259
Qy 75 LeuProGlnProSerProProAlaProThrThrThrProGlySerProProAlaProValThr 94
Db 260 CTTCCTCAACCTTCCCGCTCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 316
Qy 95 ProProThr---ArgAsnProProSerValProGlyProProSerAsnProSerArg 113
Db 317 CCCTCAAAACCTCGAAGCCCTCCATCTCTCTAACCAGGACCAACCAACATCTCCCTCA 373
Qy 114 GluGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131
Db 374 --GGATCTACTCTAGAACTCCATCAACAACTAAACCGTCCGCGCTCTGATCTCTCC 430
Qy 132 AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyValAlaLeuLeuVal 151
Db 431 GATGGATTGTCTACCGGAGTTGTGGTAGGAATCCCATCTGGTCCCAAGCTGGAGGACCTTAC 490
Qy 152 IleValThrIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db 491 ATACTGACTCTGATTGTCTCTCTGTAAGAAAGAAACGAAGAAGACACGACGATGAA 550
Qy 170 AspAlaTyrThrValProProProProProGlyProProGlyProProGlyProProGly 189
Db 551 GCTGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 610
Qy 190 GlyGlnGlnGlnThrProArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208
Db 611 GGTCAACACAGTATTGGCAACAAACAAACGCGTCACGCGCTCAGATTAATCATGTAGTG 670

```

```

209 ThrSerLeuProProProProProProProProProProProProProProProProPro 228
Db 671 ACGTCAATG---CCACACCACTAGCCTCCATCTCCACCAAGAAACCTCTCTCCGCCACT 727
Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247
Db 728 CCACCAACAGCATTCATGAGTAGCAGTGGTGTCTGATCTATTCGGATCTCTCCGGTCTT 787
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyGluGlu 267
Db 788 CCTCCACCATCTCCAGGCTTGTGTAGGCTTTCTTAAAGCACTTTCACATTATGAGGAG 847
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287
Db 848 TTGTCGAGAGCTACTAATGCTCTCTGAGGCTAATTTGTTAGCAAGGAGGGTTTGT 907
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 908 TATGTGCATAAAGTATATTGCTTAGTGGGAAAGAGTTGCTGTGAAACAGTTGAAAGCT 967
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLeuSerArgValHis 327
Db 968 GGTAGTGGTCAGGAGAGAGAGAGTTTCAGGCTGAGGTGAGATCATTTAGCAGAGTTTCA 1027
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db 1028 CACAGGCAATTTGGTCTCTTATTTGTTATGATGCGCGGTGTTCAAGAGTACTTATGTC 1087
Qy 348 TyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
Db 1088 TATGAGTTGTTTCCAAACAACATCTTGTAGTTTCACTCCATGTTAAGGGACGCGCTACG 1147
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLysSerTyrlu 387
Db 1148 ATGGAATGGAGTACTAGATTGAAGATGCTCTTGGATCTGCTAAAGGACTTTCATATCTT 1207
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407
Db 1208 CATGAAGATTGCAATCCGAAATCATTCACCGTGATTAAGGGCTCAACATATTGAT 1267
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1268 GATTTCAAATTTGAAGCTAAGTTGCTGACTTTGGTCTTGCACAAATGCTCTCTGATACA 1327
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1328 AACACTCATGTATCTACACGCTGATGGAACCTTTGGGTATTTGGCTCCGGAATATGCT 1387
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1388 GCAAGTGAAGGCTCACAGAAAGCTGACGTTTCTCATTTGGCGTTGTACTTTTGGAA 1447
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487
Db 1448 CTTATTACTGGAGGCGCCCTGTTGATGCGAACATGCTATGTATGATGACAGCTTAGTT 1507
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507
Db 1508 GACTGGGACGACCATTTGCTTAACCGGACATCTGAGGAAGAGATTTTGAAGGGTTGGCT 1567
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1568 GATTCAAGATGGGTAAATGATGATGACAGAGAGAGATGCTCGCATGTTGTTGCTTGGCT 1627
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgProArgMetSerGlnIleValArgAla 547
Db 1628 GCGCTTGTGTTGCGCATTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTAGCGGCG 1687
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1688 TTAGAAGAAATGATCGCTGCTGTGATCTTTAAAGAGGAGATGAGACCGGGTCAACGCAAC 1747

```









Db 746 ACGTCATTG---CCACCACCTAAGCCTCCATCTCCACCACGAAACCTCTCCGCACT 802

Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247

Db 803 CCACCACGACATTCATGAGTAGCAGTGGTGGTCTGACTATTCCGATCTTCCGGTCTT 862

Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267

Db 863 CCTCCACCATCTCCAGGGCTTGTGTAGGCTTTTCTAAAGACATCTTTCACATTATGAGGAG 922

Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287

Db 923 TTGTCGAGAGCTACTAATGGCTTCTCTGAGGCTAATTGTTAGGACAGGAGGGTTGGT 982

Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307

Db 983 TATGTGCATAAAGGTATATTCCTAGTGGGAAAGAGTTGCTGTGAAACAGTTGAAAGCT 1042

Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHis 327

Db 1043 GGTAGTGGTCAGGAGAGAGAGATTTTCAGGCTCAGGTTGAGATCAATTAGCAGAGTTTCAT 1102

Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347

Db 1103 CACAGGCATTTGGTTTCTTTATTGGTTATTGTTGGCCGGTGTCAAAGATTACTTGTCT 1162

Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlyArgProThr 367

Db 1163 TATGAGTTGTTCCAAACAACAATCTTGAGTTTCACCTCCATGGTAAAGGGACGCTACG 1222

Qy 368 MetGluTrpSerThrArgLeuValIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387

Db 1223 ATGGAATGGAGTACTAGATTGAAGATTGCTCTGTGATCTGTAAAGGACTTTCATATCTT 1282

Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407

Db 1283 CATGAAGATTGCAATCCGAAATCATTCACCGTATATTAAAGGGGTCAAACATATTGATT 1342

Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427

Db 1343 GATTTCAAATTTGAAGCTAAGGTTGCTGACTTTGGTCTTGGCAAGATTGCTTCGTATACA 1402

Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447

Db 1403 AACACTCATGTATCTACACGGGTGATGGAAACCTTTGGGTATTTGGCTCCGGAATATGCT 1462

Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467

Db 1463 CCAAGTGGAAAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGGA 1522

Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487

Db 1523 CTTATTACTGGAGGGCGCCCTGTTGATCGCAACAATGCTATGTAGTAGACAGCTTAGTT 1582

Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507

Db 1583 GACTGGGCACGACCATTCCTTAACCGACATCTCGAGGAAGGAGATTTTTCAGGGGTTGGCT 1642

Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527

Db 1643 GATTCAAAGATTGGTATAGTATGACAGAGGAGGATGGCTCGCATGTTGTTGGCT 1702

Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgProArgMetSerGlnIleValArgAla 547

Db 1703 GCGGCTGTGTTGCGCAATTCAGTTCGCGCAGACCTCGCATGAGCCAGATAGTACGGGCG 1762

Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567

Db 1763 TTAGAAGGAAATGTATCGCTGCTGATCTTAACGAAGGGGATGAGACCGGGTCACAGCAAC 1822

Qy 568 ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587

Db 1823 GTATACAGCTCATATGGGGAAGCACACACTATGACAGGCGCAATACACAGCAGCATG 1882

Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGlyTyrPheAlaThrGlyGlySer 607

Db 1883 ATAAAGTTTATAGGAAATGGCTCTTTGGAACTCAAGAATACGGCACACCGCGAGTACAGT 1942

Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlyGlnThrThr 627

Db 1943 AATCCAACCATGTAACGACTGTACCGTCTGGTTCAGCAGTGAAGTCAAGCCACA 2002

Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647

Db 2003 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGTCAAGGTTATAGTGGACCTCTCTT 2062

## RESULT 13

AY089024 LOCUS 2324 bp mRNA linear PLN 14-APR-2003

DEFINITION Arabidopsis thaliana clone 17909 mRNA, complete sequence.

## ACCESSION

AY089024

## VERSION

AY089024.1 GI:21407798

## KEYWORDS

Full cDNA.

## SOURCE

Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana

## REFERENCE

1 (bases 1 to 2324)

Haas,B.J., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,

Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.

Full-length messenger RNA sequences greatly improve genome

annotation

Genome Biol. 3 (6), RESEARCH0029 (2002)

22088475

PUBMED 12093376

2 (bases 1 to 2324)

Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and

Feldmann,K.

Full-length cDNA from Arabidopsis thaliana

Unpublished

3 (bases 1 to 2324)

Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and

Feldmann,K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made

available to TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of

the clones are estimated to be 5'-truncated; less than one percent

are 3'-truncated; approximately two percent represent alternative

splice variants, including unspliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be

5'-truncated if it lacks the translation initiation start (ATG). A

sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA

sequences are derived from the Ws or Laer ecotypes and therefore

may contain polymorphisms when compared to sequences from Col-0.

Genet carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the

5' sequences, selection of clones, and sequence assembly.

## FEATURES

Location/Qualifiers

1..2324

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/clone="17909"

## ORIGIN

Alignment Scores:

Pred. No.: 2,45e-62 Length: 2324

Score: 2885.00 Matches: 570

Percent Similarity: 90.30% Conservative: 26

Best Local Similarity: 86.36% Mismatches: 42

Query Match:	83.55%	Indels:	23
DB:	8	Gaps:	13
US-10-086-464-2 (1-647) x AY089024 (1-2324)			
Qy	1	MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer	17
Db	105	ATGTCACAGCGCGCTCTCCAGGCACTACTCCATCACCACCTCCACCGCTCTCCCTCCACCA	164
Qy	18	AsnSerThrThrThrProProAla-----SerAlaProProProThrPro	35
Db	165	AACTCGACACACACACTCTCTCCAGCAGCTCTTCTCTCTCTCCACACCACTCTCT	224
Qy	36	SerSerProProPro-----ProSerThrIleProThrSerProProProSerSerArgSer	54
Db	225	TCTTCTCTCTCTCGCGCGTCAACTAATTCACACCTCTCTCTCTCTCTCTCTCTCTCTCT	284
Qy	55	ThrProSerAlaProProProSerProProThrProThrProSerThrProGlySerProProPro	74
Db	285	CCT	332
Qy	75	LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr	94
Db	333	CTTCTCTCAACCTTCTCCATCCGCTCCCATCACTCTCTCTCTCTCTCTCTCTCTCTCTCT	389
Qy	95	ProProThr-----ArgAsnProProProSerValProGlyProProProSerAsnProSerArg	113
Db	390	CCCTCAAAACCTCGNAGCCCTCCATCTCTTAACCAAGGACCAACCAACACTCTCTCTCTCT	446
Qy	114	GluGlySerProArgProProSerSer-----ProSerProProProSerProSerSer	131
Db	447	---GGATCTACTCTCTAGAACTCCATCAAAACGCTCAAAACGCTCGCCACCGCTCTGATTTCT	503
Qy	132	AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyValAlaLeuVal	151
Db	504	GATGGATTGTCTACCGGAGTTGTGGTAGGAATCGCCATTGGAGGAGTCTGATTTCTTGT	563
Qy	152	IleValThrLeuIleCysLeuLeuCysIleValValValValValValValValValVal	169
Db	564	ATACTGACTCTGATTTGTCTCTCTGTAAGAAAGAAACGAAAGAAAGACAGACAGATGAA	623
Qy	170	AspAlaTyrrValProProProProProProProGlyProGlyProGlyProGlyProGly	189
Db	624	GCTGCTTACTATGTT---CCT	680
Qy	190	GlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	208
Db	681	GGTCAACAACAGTATTGGCAACCAACAAACCGCTCAACCGCGCTCAGATAATCATGTAGTG	740
Qy	209	ThrSerLeuProProProProProProProSerProProProProProProProProPro	228
Db	741	ACGTCAATG---CACACACTAAGCTTCATCTCTCAACCAAGAAACCTCTCTCTCTCTCT	797
Qy	229	ProProPro-----PheMetSerSerSerGlySerAspTyrrSerAspArgProValLeu	247
Db	798	CCACCACAGCATTCATGAGTAGCAGTGGTGGTCTGACTATTCTGATCTCTCTCTCTCTCT	857
Qy	248	ProProProSerProGlyLeuValLeuGlyPheSerIleSerIleSerIleSerIleSerIle	267
Db	858	CCTCCACCATCTCCAGGCTTGTGTAGGCTTTTCAAAAAGCACTTTCACTTATGAGGAG	917
Qy	268	LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly	287
Db	918	TTGCAAGAGCTACTAATGGCTTCTCTGAGGCTAATTTGTTAGGACAGGAGGTTTGGT	977
Qy	288	TyrValHisIleValLeuProSerGlyValValAlaValIleValIleValIleValIle	307
Db	978	TATGTGCATAAGGTATATTCCTTAGTGGGAAAGAGTTGCTGTGAAACAGTTGAAAGCT	1037
Qy	308	GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis	327
Db	1038	GGGAGTGTCTAGGAGAGAGAGAGTTTTCAGGCTCAGGTTGAGATCATTAGCAGAGTTTCA	1097
Qy	328	HisArgHisLeuValSerLeuValGlyTyrrCysIleAlaGlyAlaIleAlaIleAlaIle	347
Db	1098	CACAGGCAATTTGGTTTCTCTTATTGGTTATTATGATGCTGATGTTCAAGAGATTCTTGT	1157
Qy	348	TyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr	367
Db	1158	TATGAGTTTCTTCTCTCAACAATCTTGGAGTTTCCACCTCCATGTTAAGGAGCCCTACA	1217
Qy	368	MetGluTrpSerThrArgLeuValIleAlaLeuGlySerAlaIleGlyLeuSerTyrrLeu	387
Db	1218	ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGGATCTGCTAAGGACTTTCATATCTT	1277
Qy	388	HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle	407
Db	1278	CATGAAGATTGCAATCGAAATCATTCACCGTGATATTAAAGGCTCAACATATTGATT	1337
Qy	408	AspPheIlePheGluAlaIleValAlaAspPheGlyLeuAlaIleIleAlaSerAspThr	427
Db	1338	GATTTCAAATTTGAAGCTAAGTTGCTGACTTTGGTCTTCCAGGATTGCTTCTGATACA	1397
Qy	428	AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrrLeuAlaProGluTyrrAla	447
Db	1398	ACACTCATGTATCTACACCGGTGATGGAAACCTTTGGGTATTGCTCCAGATATGCT	1457
Qy	448	AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu	467
Db	1458	GCAAGCGGAAAGCTCACAGAAAAGTCTGACGTTTCTCATTTGGTGTGTTGCTTTGGAG	1517
Qy	468	LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrrValAspAspSerLeuVal	487
Db	1518	CTTATAACCGGAGGCGCCCTGTTGATGCAACAATGCTATGTAGATGACAGCTTAGTT	1577
Qy	488	AspTrpAlaArgProLeuLeuAsnAsnAlaSerGluGlnGlyAspPheGlyGlyLeuAla	507
Db	1578	GACTGGCAGACGCTTCTGCTTAAACCGCATCTGAGGAGAGAGATTTTGATGTTGGCT	1637
Qy	508	AspAlaLysMetAsnAsnGlyTyrrAspArgGluGluMetAlaArgMetValAlaCysAla	527
Db	1638	GATTCAAAGATGGTAATGAGTATAG-AGAGAGGAGATGCTCCCATGTTGCTTGGCT	1696
Qy	528	AlaAlaCysValArgHisSerAlaArgArgProArgProArgMetSerGlnIleValArgAla	547
Db	1697	GCAGCTTGTGTTGCCATTCAGCTCGCCGACAGCTCGCATGAGCCAGATAGTACGGCG	1756
Qy	548	LeuGluGluValSerLeuSerLeuAsnGluGlyMetArgProGlyGlnSerAsn	567
Db	1757	TTAAGAGAAATGTATCGCTGCTGATCTTAAAGAGGATGAGACCGGGTCAACGCAAC	1816
Qy	568	ValTyrrSerSerTyrrGlySerThrAspTyrrAspSerSerGlnTyrrAsnGluAspMet	587
Db	1817	GTATACAGCTCATATGGAGGAGCACTGACTATGACACGAGCAATACACGATGACATG	1876
Qy	588	LysIlePheArgLysMetAlaLeuGlyThrGlnGluTyrrAsnAlaThrGlyGlyTyrrSer	607
Db	1877	AAGAAGTTTAGGAAATGGCTCTTGGAACTCAAGAATACGGCAACACCGCGAGTACAGT	1936
Qy	608	AsnProThrSerAspTyrrGlyLeuTyrrProSerGlySerSerGlyGlnThrThr	627
Db	1937	ATCCCAACCACTGACTACGAGCTGTACCCCTCTGGTCTCAAGCAGTGAAGTCAAGCCACA	1996
Qy	628	ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrrSerGlyProSerLeu	647
Db	1997	CGAGAAATGGAGATGGGAAAGATTAGAAAACCGGTCAAGGTTATAGTGGACCTCTCTT	2056
RESULT 14			
AB020746			
LOCUS			
DEFINITION			
Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MOB24.			
ACCESSION			
AB020746.1 GI:3985949			
VERSION			
AB020746.1			
KEYWORDS			
Arabidopsis thaliana (thale' cress)			
SOURCE			
Arabidopsis thaliana			

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites)

REFERENCE  
AUTHORS  
TITLE  
Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S. Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC and BAC clones  
DNA Res. 7 (3), 217-221 (2000)

JOURNAL

MEDLINE

PUBMED

20363099

REFERENCE

10907853

AUTHORS

2 (bases 1 to 79706)

Kaneko,T., Kato,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.

Direct Submission

Submitted (03-DEC-1998)

Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,

Tel:81-438-52-3935, Fax:81-438-52-3934)

Address for correspondence: kaos@kazusa.or.jp

For the latest information on annotation of this clone, please see

http://www.kazusa.or.jp/kaos/cgi-bin/aggd.graph.cgi?c=MOB24

Genes with similarity to proteins in the databases are described in

'product' or 'note' qualifiers. Genes that have no significant

protein similarity are described as 'unknown protein'.

The software programs used to predict genes include: Grail

(Informatics Group, Oak Ridge National Laboratory,

http://compbio.ornl.gov/Grail-1.3/),

GENSCAN (Chris Burge, MIT, http://CCR-081.mt.edu/GENSCAN.html),

NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of

Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and

SplicePredictor (Volker Brendel, Stanford University,

http://grainlin.zool.iastate.edu/cgi-bin/sp.cgi).

Genes encoding tRNAs are predicted by tRNAscan-SE

(Sean Eddy, Washington University School of Medicine, St. Louis,

http://genome.wustl.edu/eddy/tRNAscan-SE/).

This sequence may not be the entire insert of this clone. It may be

shorter because we remove overlaps between neighboring submissions.

The 5' clone is MXP5 and the 3' clone is MSD24.

Location/Qualifiers

1..79706

/organism="Arabidopsis thaliana"

/mol\_type="genomic DNA"

/strain="Columbia"

/db\_xref="taxon:3702"

/chromosome="3"

/clone\_lib="Mitsui P1"

/clone="MOB24"

2398..3399

/note="gene id:MOB24.1

unknown protein"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="BAB01995.1"

/db\_xref="GI:9294038"

/translation="MGSEDETGYKPKFYSLNRONHPMYSRPIPKRAYNNEDEDEB

VEGEEDPQGYIIGNERFQRKPKNVSVFAGSDAKVADWREQEAFLVLEW

GDRLQGRSLRNEDMNEAEKVSLEURMEKSETQCRMTDDDKRKREKI KVEKS

GLSSKWSFFNKJMLLCKSPKSLGLACGVDSGEFVPMNTKTVLYDKSNGFDMNDSP

GDSEEDDEVEYRKVNDAAASYKMLADSVRFKGYEKMEKSKKEQMKLEKMR

ADFQDLLEKQKQIVDRAQSEIARLREEENHHGGDDDEDEBEMENDSDVNLSD"

complement (join (3794..4276, 4363..4490, 4577..4673,

4757..5179, 5354..5548, 5693..5970, 6088..6214, 6297..6496,

6601..6648, 6743..6776, 6882..7013, 7464..7559, 7677..7743,

8162..8217, 8290..8385, 8470..8998, 9122..9462))

/note="gene id:MOB24.2"

/codon\_start=1

/evidence=not\_experimental

/product="DNA repair protein MutS"

/protein\_id="BAB01996.1"

/db\_xref="GI:9294039"

/translation="MQQRSLISFFQKFQFTATTKGLVSGDAASGGGGGGPRFNKRG

DAKGDASVRFAVSKSVDEVRGTDTPPEKVPVRVLPSPGPKPAESAGDASSLPFNMHKF

VKVDKDCSGERSREDVVPFLNDSLLCMKANDVIFQFRSNNKTKQERNHAFSFGRAEL

RSVEDIGVDGDPGPETPGMRPRASRLKRVLEDEMTFKEDKVPVLDSNKKRLKMLQDPV

CGKXKNEGTKEWLESSIRUDANRRPDDPLYDKRTLHI PPDVEKMSASOKQVWS

QKSEAYMDLVFFKVGKFIYELDAELGHELDKWMKMSGVGKCRQVIGSISGIDEAV

OKLLEAGYKVRIEQLETSQAKARGANTIIPRKLVOVLFPSTASEGNIGPDVHLLA

IKETIMELQKCVTYGFAFVCAALRFVGSISDDASCAALGALLMQVSPKEVLYDSK

IGLSREAKLRKTYLTGSTAVQLAPVQVMDTDAAGVRNI IBSNGYFKGSSSWNCA

VDGLNECDVALSALGELINHLRLKLELDVLKHGDIIPPQVYRGCLRIDGQTMVNLTF

NSCDGSPSTLYKYLDCNVSPGKLLRNWI CHPLKDVESINKRLDVEEFANSES

MQITGYHLKLPDLERLLGRIKSVSRSSASVLPALLQKKVLKQKVAFGQIVKGFSG

IDLLALQKESNMWSLYLKCLKPILVKGSGLELFSQFAAIDSDFPNTQVDTDE

NAETLTILIELFIERTATQWSEVIHTISCLDTRSFALFAASLSAGSMARPIFPESAE

DQNKTKGPIKIQGLWHPFAVAGQLPVPNDLILGEARRSGSIHPRSLILTGNM

KGSTLLRATCLAVIPAOLGCVYPCSCETSLVDITFTRLCASDRIMTGESTFLVECT

ETASVLQATQDSILVILDELGRGTSTFDGYAIASVFRHLVEKVCQCHMLFATHVHLIT

KFASHPRTVSKMACAFKSRSDYQPRGCDQDLVFLRLTEGACSPGSLQVALMAGI

PNQVETASGAQAQMKRSGIGENFKSELSESEFSLSHEDWLKSLVGISRVAHNNAPIGE

DDYDTLFLMHEIKSSYCVPK"

10131..10577

/note="gene id:MOB24.3"

/codon\_start=1

/evidence=not\_experimental

/product="ethylene-responsive transcriptional

coactivator-like protein"

/protein\_id="BAB01997.1"

/db\_xref="GI:9294040"

/translation="MPSRYPGAVTDQWEPVVLHKSQKODLRDPKAVNALRNGVAV

QTVKFDAGSNKKGKSTAVFVINTKLEESTEPAMDVRVAEVRMLQKARLEKMSQ

ADIAKQINERTQVQVEYKQVAPNQVAPQAKMEKVLGVKLRGKIGK"

complement (join (11108..11381, 11467..11528, 11699..11836,

11934..12071, 12161..12334, 12421..12510, 12718..13104,

13623..13765, 14306..14405))

/note="gene id:MOB24.4"

/codon\_start=1

/evidence=not\_experimental

/product="aldehyde dehydrogenase"

/protein\_id="BAB01998.1"

/db\_xref="GI:9294041"

/translation="MENGKCNGATTVKLPEIKETKLFINGQFIDAASGKTFETIDPN

GEVITAIEGKEDVDLVANVAFYDFHGPWPMTGFERAKLINKFADLLEENIELA

KLDVADGGKLFQLGKADIATAGHFRYNAGAADKHGETLKMTRQSLFGYTLKEPIG

VVGNIIIPWNPFSIMFATKVPAMAACTVMPAEQTSLSALFYAHLKSEAGIPDGLV

NIVTFGSGTAGAAIAASHMDVKVSTGSTDVGRKIMQAAAANLKKVSLBELGKSPLL

IFNDADIDKAADLALGCFYNKGEICVASSRVFQEGIYDKVVEKLVKAKDWTVGDP

FDSTARQGPQVDKROFEKILSYIEHKNEGATILLTGKAI GDKGYFIQPTIEADVTED

MKYIQDEIFGPMVKMKFTVEGICKANNTKYGLAAGILSQDIDLINTVSRSIKAGI

IWNCYGFDDLDCPYGSKMNGNCRSGMDALDNLQTKSVVMPLHNSPMW"

complement (join (19742..19933, 20400..20657))

/note="gb|AAD25142.1

gene id:MOB24.5

similar to unknown protein"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="BAB01999.1"

/db\_xref="GI:9294042"

/translation="MTPLSISASGSPWPLLSPPPLSSSLNRAAPISFSPFSRINHGF

SCSFIQVORALSRTKQIQVVAMAPEEKLTRNLPDFIEWERPKGRRPDIFFK

FSPKMSPLPPMPDYDPPEEEBEEKEETEDDPEKEDDQDNRS"

join (23402..23459, 23819..24021)

/note="gene id:MOB24.6

unknown protein"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="BAB02000.1"

/db\_xref="GI:9294043"

/translation="MKNASLKPLLLFILVTSNLAGAERKLTGVADVIVEGEAASPO

YGIIDENDTLNKSQSKCRDIDCTFQCRRGFCNLILQKCECL"

complement (join (25945..26640, 26714..26824, 26914..27012,

27119..27224, 27291..27360, 27436..27583))

/note="contains similarity to ubiquitin conjugating enzyme

gene id:MOB24.7"

/codon\_start=1

/evidence=not\_experimental

/protein\_id="BAB02001.1"

CDS

CDS

CDS

CDS

CDS

```
/db_xref="GI:9294044"
/translation="MAQARLSLRMOKELKLLSDPPHGHASFPHLSSAANGSGDFSTP
STIDAOLEGPDYVANGIENLKIQIPERVFPQPIVSEATPIVHPNIDNSGRICLDI
LNLPPKANQWSPSLNISTVLSMRLLSEPNPDDGLMCEVSREYKYNQTFDYKAREMT
EDNGKRAVGVGEGNSFGNDGIKRKKSLALPSQKKOIKGEBEELTRGVSAAC
KENPKNLNGKKLSGLKQPCNDTLASPTSAKSDNNRSLRKLRLSPGLGELNEVSK
PEVLAQTDMLKMNQNDARSURGEFENSVLEETSMAESIVLSDSDSGQEBEERVS
SRSLSLAKRRLVLCRP"
```

CDS

```
28636..30168
/note="emb|CAB79678.1
gene id:MOB24.8
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein id="BAB02002.1"
/db_xref="GI:9294045"
```

```
/translation="MSSLIAVLSFNSFPLKLSRILSNLMLLVRTSTNKKRWCL
AIPYHLHATPLWKSSTIMLLRFQSMVEPTTYFEISLRAFILQAMSHGETLLWD
TTSFLGLTRILWFSASPHGQOQTSTFPSPENDASLSSPVNSRRHHDLHSEGLK
GLQVLGLVFGPLAYRMPISKDDPKLPLKPRSDLYSLISFOSLVHLLASIDRHH
TNNELVLIHLKNSLPTLPLVSRFTVTGFICLKNGIMPEPTTSRWCDLCHPSATI
CISPTHLGISFKQSCISRVKAHISRSKEVPNLFELITEPSHALVWKGKIGIMI
TSPSGDYQRFNLHSHSYPLTEPKQRKNLIFEIHDLOSQKNGIMISPPWSDYR
SFFNLDLPTIVTHLGDQQLLVLSARTSALEPYSTSRLLTVTNLLSIDSFVED
PSTTRDLTCAIMPYSMLKALMFMFSFHIIIVHLVGNASFNVYVNFESLLLIFF
LI"
```

CDS

```
join(32831..33058,33143..33907)
/note="contains similarity to heat shock transcription
factor"
```

```
gene id:MOB24.9"
/codon_start=1
/evidence=not experimental
/protein id="BAB02003.1"
/db_xref="GI:9294046"
/translation="MEDDNNNNNNNNVIAPIVKTQVMNDPSTDWLITWGPANHNSF
IVDPLDSQRIILPFAKNNFSSVFLQNTYFRKVDPPDWEFANFHLRQKHLLN
NIARKKARGMYGQDGEIVREIERLKEORELEABIORNNRRIEATEKRPQOMVA
FLYKVRDPLLPMWLEKERTKQOQVSDKKERVTMTSVKSEEREVEDEGRFVRV
MSSSTPSPSTENLYRHSFDGVLVPMIQQFGSYETGLVAKSNLSNSTSTSSSLTS
TFSLPESVGGGGGCGSIOGERRYKETATFGVVENPPTTPPYPSLFRGGF"
completion (join(37054..37590,37670..37809,37894..38197,
38288..38398,38486..38584,38662..38772,39019..39066,
39311..39367))
/note="gene id:MOB24.10
unknown protein"
/codon_start=1
```

## Alignment Scores:

Pred. No.:	7,29e-55	Length:	79706
Score:	2622.00	Matches:	573
Percent Similarity:	61.19%	Conservative:	23
Best Local Similarity:	58.83%	Mismatches:	43
Query Match:	75.93%	Indels:	337
DB:	8	Gaps:	19

US-10-086-464-2 (1-647) x AB020746 (1-79706)

Qy	1	MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSerProSer 17
Db	51787	ATGTCCACAGCGCGTCTCCAGGCACTACTCCATCCACCATCTCCACCGCTCTCTCCACCA 51846
Qy	18	AsnSerThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db	51847	AACTCGACCAACCACTCTCTCTCCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 51906
Qy	36	SerSerProProPro-----ProSerThrIleProThrSerProProProSerSerArgSer 54
Db	51907	TCTTCTCTCTCTCGTCGCGCGTCAACTAAATCAACCTCTCTCTCTCTCTCTCTCTCTCTCT 51966
Qy	55	ThrProSerAlaProProProSerProProThrProThrProSerThrProGlySerProPro 74
Db	51967	CCTCCTTCT 52014

Qy	75	LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
Db	52015	CTTCTCTCAACTTCCCTCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 52071
Qy	95	ProProThr---ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
Db	52072	CCCTCAAAACCTCGAAGCCCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 52128
Qy	114	GluGlyGlySerProArgProProSerSer-----ProSerProProProProSerSer 131
Db	52129	---GGATCTACTCTAGAACCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACT 52185
Qy	132	AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyValAlaLeuLeuVal 151
Db	52186	GATGGATGTCTACCGGAGTTGTGTAGGAATGCCCATTCGGAGGAGTCGTATTCTTGT 52245
Qy	152	IleValThrLeuLeuCysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db	52246	ATACTGACTCTGATTTCT 52305
Qy	170	AspAlaTyrrTyrrValProProProProProProProProProProProProProProPro 184
Db	52306	GCTGCTTACTATGTTCT 52364
Qy	184	----- 184
Db	52365	CAACTTCATGAATAAATAAATCGAATATTAACCATAGTTGTTCAATTCTGATACGATTTG 52424
Qy	184	----- 184
Db	52425	CTCAAACTCTGATCTACTTGTATTGATTGATTGTTTATCTGTTAAAGAGATTACCTTTGTGGA 52484
Qy	185	-----GlyGlyProTyrrGlyGlyGln 191
Db	52485	GTTCAAACTTTACAGAGATTGATGTTGTTTATCTAGCTCGAGGACCTTACGGTGGTCAA 52544
Qy	192	GlnGlnGlnTrpArgGlnGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 210
Db	52545	CAACAGTATTGGCAACAAACAAACGCGTCACGCGCGTCAGATATCATGTAGTAGACGTCA 52604
Qy	211	LeuProProProProProAlaProSerProProArgGlnProProProProProProProPro 230
Db	52605	TTG---CCACCACCTAAGCCTCCATCTCCACCAAGAAACCTCTCTCGCCACCTCCACCA 52661
Qy	231	Pro---PheMetSerSerSerGlyGlySerAspTyrrSerAspArgProValLeuProPro 249
Db	52662	CCAGCATTCATAGTAGACGTGGTGGTCTTGACTATTTCGGATCTTCCGGTTCTTCTCTCCA 52721
Qy	250	ProSerProGlyLeuValLeuLeuGlyPheSerLysSerThrPheThrTyrrGluGluLeuAla 269
Db	52722	CCATCTCCAGGCGTTGTGTAGGCTTTTCTTAAAGCACTTTCTATATGAGGAGTTGTCG 52781
Qy	270	ArgAlaThrAsnGlyPheSerGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 289
Db	52782	AGAGCTACTAATGGCTTCTCTGAGGCTAATTTGTAGGCAAGGAGGTTTGGTTATGTG 52841
Qy	290	HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
Db	52842	CATAAAGGTATATTTGCCCTAGTGGGAAAGAGTTGCTGTGAAACAGTTGAAAGCTGGTAGT 52901
Qy	310	GlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLeuSerArgValHisHisArg 329
Db	52902	GGTCAGGAGAGAGAGAGTTTTCAGGCTGAGTTGAGATCATATAGCAGAGTTTCATCAGG 52961
Qy	330	HisLeuValSerLeuValGlyTyrrCysIleAlaGlyAlaLysArgLeuLeuValTyrrGlu 349
Db	52962	CATTGGTTTCTCTTATTGTTTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 53021
Qy	350	PheValProAsnAsnAsnLeuLeuLeuHisLeuHisGlyGlu----- 363
Db	53022	TTTGTTCCAAACCAACATCTTGAGTTTTCACCTCCATCGTTTA-GAACAACAACTAACTCTT 53080

QY	363	-----	363
Db	53081	TCCTTTGANTTTGTTACATTCGCTTTGTTGTTTTCAGTTAGTCATGATTTGTTGGGAATT	53140
QY	364	-----GlyArgProThrMetGluTrpSerThrArgLeuLysIleAla	377
Db	53141	GAATTTGTTATAGTTAAGGACGGCTACGATGGNATGGATCTAGATTTGAAGATTGCT	53200
QY	378	LeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys-----	391
Db	53201	CTTGATCTGCTAAAGGACTTTTCATATCTTCATGAAGATTGTAAGTCTFAAAACCTTAAAA	53260
QY	391	-----	391
Db	53261	ATTATGTACAGATGTTTGATAAATACTGCTCTTTGTTGTCCTGAGGTPATATATAAAT	53320
QY	392	-----AenProLysIleIleHisArgAspIleLysAlaSerAenIleLeuIleAs	408
Db	53321	ATTTCGAGCGAATCCGAANAATCATTCACCGTGATATTAAGCGGTCAAAACATATTGATTTGA	53380
QY	408	pPheLysPheGluAlaLys-----	414
Db	53381	TTTTCAAATTTGAAGCTAAGGTATGTTTACTCGATATGTTTGTGAAGTGTCTTTTGTAAAT	53440
QY	415	-----ValAlaAspPheGlyLeuAlaLysIleAlaAs	425
Db	53441	TGTGTCACTAAAGCTCTTTCTTTTCAGGTTGCTGACTTTTGGTCTTGCCGAAGATTGCTT	53500
QY	425	erAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly-----	440
Db	53501	CTGATACAAACACTCATGTATCTACACCGGTGATGGGAACCTTTGGGTAAAAATCACTTT	53560
QY	440	-----	440
Db	53561	TAATCTACTAAACATGTACATTTATGTCAGTTGCTTTAATGTCATGTGATCTTGTGCAACA	53620
QY	441	-----TyrLeuAlaProGluTyrAlaAlaSerG	450
Db	53621	AGTCTGAAACTTTTGGTTTGGGATTTTAGGTATTTGCGTCCGGAATATGCTGCAAGTG	53680
QY	450	lyLeuLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleT	470
Db	53681	GAAGCTCACAGAAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGGAACTTATTA	53740
QY	470	hrGlyArgProValAspAlaAenAenValTyrValAspAspSerLeuValAspTrp-	489
Db	53741	CTGGAGGCGCCTGTTGATGCGAACATGCTATGTAGATGACAGCTTAGTTGACTGGG	53800
QY	489	-----	489
Db	53801	TACTTTACTATGCTGCAGATAAATCTAAACTTCTGATTTTCTACAAATTTAACTGTTGAC	53860
QY	490	-----AlaArgProLeuLeuAenArgAla	497
Db	53861	TAAACAAAAGCTATGCTGTGATTGATGTGTGGAACGGCACCACTTCTTTAACCGAGCA	53920
QY	498	SerGluGlnGlyAspPheGlyLeuAlaAspAlaLysMetAenAsnGlyTyrAspArg	517
Db	53921	TCTGAGGAAGGAGATTTTGGGGTTTGGCTGATTCAAAGATGGGTAAATGAGTATGACAGA	53980
QY	518	GluGluMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArg	537
Db	53981	GAGGAGATGGCTCGCATGTTGCTTGGCGTGGCGTCTGTTGTCGCCATTCAGTCGCGCG	54040
QY	538	ArgProArgMetSer-----	542
Db	54041	AGACCTCGCATGAGCCAGGTAACTTAATCCTTTCTCTAGAGAGACATATTGTTGG	54100
QY	542	-----	542
Db	54101	TTCTGTTTATTGGTGAATAATATGTTTCCGTATGTTTAAAGAAAGGTTTGTGATG	54160
QY	542	-----	542

Db	54161	GTTTTGTGAGAAAGATCGGTTTCAGTTATGTTCTTATGCTCCTAGCCCTACTCTTTTGTACC	54220
QY	542	-----	542
Db	54221	CATCATCATACTATCTAAACAACACCCGATTGAATATGTCAAAAGAAATTTGGCTGCAGTTTC	54280
QY	542	-----	542
Db	54281	TTTTGTTTTCATTTTGTCCAAAGGATTTGTCAATTTTAAACATATGTTTAATGCTAATGGAGTGAG	54340
QY	543	-----GlnIleValArgAlaLeuGluGlyAenValSerLeuSer	555
Db	54341	TGATTTGCTGTACAAAATAGCAGATAGTAGCGGCGTTAGAGGAATATGATCCTGTCT	54400
QY	556	AspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySer	575
Db	54401	GATCTTAAACGAAGGATGAGACCGGGTCACAGCAACGTATACAGCTCATATGGAGGAAGC	54460
QY	576	ThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeu	595
Db	54461	ACAGACTATGACGACGACCAATACACACGACGACATGATAAAGTTTAGGAAATGGCTCTT	54520
QY	596	GlyThrGlnGluTyrAenAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeu	615
Db	54521	GGAACTCAAGAATACGCCACACACCGCGAGTACAGTAATCAACACGTAGTACCGACTG	54580
QY	616	TyrProSerGlySerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIle	635
Db	54581	TACCCGCTCGTTCAACGACGTGAAGGTCAAGCCACACGAGAAATGGAGATGGGAAAGATT	54640
QY	636	LysArgThrGlyGlnGlyTyrSerGlyProSerLeu	647
Db	54641	AAGAAAACCGTCAAGGTTTATAGTGGACCTCTCTT	54676

RESULT 15			
LOCUS	AK103247	2307 bp	mRNA linear
DEFINITION	Oryza sativa (japonica cultivar-group) cDNA clone:J033123K23, full insert sequence.		PLN 24-JUL-2003
ACCESSION	AK103247		
VERSION	AK103247.1	GI:32988456	
KEYWORDS	FLU_CDNA; CAP trapper.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	1		
AUTHORS	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Genome Sequencing & Analysis Group, Otonoda, Y., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oato, N., Oka, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.		
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice		
JOURNAL	Science 301 (5631), 376-379 (2003)		
MEDLINE	22752273		
PUBMED	12869764		
REFERENCE	2 (bases 1 to 2307)		
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,		



Job time : 5988 secs

```
Db      988  |||||TATGTTCAAAAGAGTTCTCGGATGGGACAGAGGTGCTGTGAACCAATTGAGAGAT 1047
Qy      308  GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327
Db     1048  GGGAGTGGGACGGAGGACGTTGAATTTTCAGCGGAGGTTGAGATTATCAGCGGGTACAT 1107
Qy     328  HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuVal 347
Db     1108  CACAAGCATCTTGTAAACATTGGTGTATTGTCATTCTCGAGGGAAGAGGTTGCTTGTG 1167
Qy     348  TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
Db     1168  TATGATGATGTTCCAAACAACACATTGGAGCTCCACTTGCATGGAGAGGCCGACCAACC 1227
Qy     368  MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrIleu 387
Db     1228  ATGGAATGGCTTACAAGATTACGCATTGCTTTGGTGTCTGCGAAGGGATTGGCATATCTT 1287
Qy     388  HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleIle 407
Db     1288  CATGAAGATTGCCATCTCTAAGATCATTCCCGTGACATAAAGTCAGCAAAATATTCTTCTT 1347
Qy     408  AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db     1348  GACGCGAGATTTGAGGCAAGGTGGCAGATTTCGGACTTGCATAAATTGACTTCTGATAAT 1407
Qy     428  AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db     1408  AACACTCATGTTTCAACCAGAGTAATGGGCACATTTGGGTACCTTGGCACAGAGTATGCG 1467
Qy     448  AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db     1468  TCTTCTGGCCAGCTAACTAGAGAAATCAGATGCTCTTTCTTCGGAGTAATGCTTCTTGAG 1527
Qy     468  LeuIleThrGlyArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db     1528  CTAATAACTGGCCGCGCTCTGTGAGATCGAACCATCACAGATGGATGACAGCTTGTT 1587
Qy     488  AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db     1588  GATTGGCAAGGCTTTAATGATGCGAGCATCAGATGATGGCAATTATGATGCTTTAGTG 1647
Qy     508  AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db     1648  GATCCACGGCTAGGACAGAGTACAAATGGTAAACGAGATGGCGAGGATGATTGCTTGCA 1707
Qy     528  AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db     1708  GCTGCTTGTTGACGCCATTCTGCACCGCAGACGCCACGGATGAGCCAGGTTGTTGGGCC 1767
Qy     548  LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db     1768  TTGGAAGCGCATGTCTGATGATCTGAACGAAGGTTTCTGCTCTGGCCATAGCCGA 1827
Qy     568  ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db     1828  TTCTTTGGGATCATAC---AACAGCAACGAGTACGACACGGGCCATTACAATGAAGACCTG 1884
Qy     588  LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db     1885  AAAAAATTCAGGAAGATGGCATTTGGCAGTGGCAATCTA-----GAGAGCAGC 1932
Qy     608  AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
Db     1933  CAGCAACCGCAGCCACCGAGTTCGTCCCGAATCGATCCGTATCAATGGGGGATGCGCGT 1992
Qy     628  Arg-----GluMetGluMetGlyLysIleLysArgThrGly 639
Db     1993  CAAATACCGGAGACGGAGATGGGAGTTTGAAGAAGATGGC 2040
```

Search completed: April 25, 2004, 00:00:59

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 24, 2004, 18:28:17 ; Search time 604 Seconds  
(without alignments)  
4550.636 Million cell updates/sec

Title: US-10-086-464-2  
Perfect score: 3453  
Sequence: 1 MSSAPSGTGTGPPSPNST.....REMEMGKIRKTGCGYSGPSL 647

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10086464/runat\_23042004\_083014\_5603/app\_query.fasta\_1.839  
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10086464 -CGN 1 1 708 -runat\_23042004\_083014\_5603 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	3453	100.0	1944	4	AAF77094	Asf77094 Arabidops
2	3453	100.0	1944	9	ACF36548	B. napus
3	3453	100.0	2189	9	ACF36549	B. napus
4	2936.5	85.0	1959	9	ACF38557	Ac38557 A. thalia
5	2885	83.6	2324	3	AAAC36968	Arabidops
6	1821.5	52.8	1812	9	ACF36554	O. sativa
7	1668	48.3	1515	9	ACF36555	Z. mays
8	1637.5	47.4	1488	9	ACF36556	G. max

9	1507	43.6	2520	4	AAF77096	Arabidops
10	1504	43.6	2520	9	ACF36551	Arabidops
11	1479.5	42.8	2820	4	AAF77097	Arabidops
12	1471.5	42.6	2820	9	ACF36552	Arabidops
13	1209.5	35.0	2880	4	AAF77095	Arabidops
14	1209.5	35.0	2880	9	ACF36550	Arabidops
15	1185	34.3	3060	9	ACF36553	Arabidops
16	1181	34.2	3060	4	AAF77098	Arabidops
17	839.5	24.3	1923	7	ADAY0890	Rice gene
18	816.5	23.6	2124	6	ABZ13274	Arabidops
19	804.5	23.3	2865	7	ADA70701	Rice gene
20	802.5	23.2	2011	4	AAH22557	P. patens
21	797	23.1	1730	7	ADA70647	Rice gene
22	793.5	23.0	2823	7	ADA69478	Rice gene
23	782.5	22.7	1419	3	AAC43029	Arabidops
24	782.5	22.7	1419	6	ABZ13121	Arabidops
25	775.5	22.5	1626	3	AAC48313	Arabidops
26	775.5	22.5	1917	7	ADA70527	Rice gene
27	773	22.4	1479	3	AAC43040	Arabidops
28	769.5	22.3	1353	3	AAC42860	Arabidops
29	769.5	22.3	1353	6	ABZ14106	Arabidops
30	766	22.2	1494	3	AAC48560	Arabidops
31	758.5	22.0	1572	3	AAC42898	Arabidops
32	758.5	22.0	1572	6	ABZ13705	Arabidops
33	758.5	22.0	1920	6	AAL45905	A thalian
34	758.5	22.0	2012	3	AAC44847	Arabidops
35	756	21.9	3239	3	AAAY9271	Eucalyptu
36	754.5	21.9	2136	7	ADA69829	Rice gene
37	749.5	21.7	1751	3	AAC45313	Arabidops
38	744.5	21.6	1485	3	AAC42702	Arabidops
39	740.5	21.4	3525	7	ADA71170	Rice gene
40	740	21.4	2868	3	AAAY9266	Pinus rad
41	740	21.4	3153	7	ADA71066	Rice gene
42	736.5	21.3	1703	3	ADA66629	Zea mays
43	736	21.3	3624	7	ADA69512	Rice gene
44	735	21.3	2004	3	AAC40780	Arabidops
45	734	21.3	1455	3	AAC45877	Arabidops

ALIGNMENTS

RESULT 1  
AAF77094  
ID AAF77094 standard; DNA; 1944 BP.  
AC AAF77094;  
XX  
XX  
DT 17-MAY-2001 (first entry)  
XX  
DE Brassica napus PERK1 DNA.  
XX  
KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.  
XX  
OS Brassica napus.  
XX  
PN WO200114563-A1.  
XX  
PD 01-MAR-2001.  
XX  
PF 18-AUG-2000; 2000WO-CA000966.  
XX  
PR 19-AUG-1999; 99US-0149466P.  
PR 13-OCT-1999; 99US-0159122P.  
XX  
PA (GORI// GORING D.  
(SILV// SILVA N.  
PI Goring D, Silva N;  
XX  
DR WPI; 2001-244305/25.  
XX  
PT New proline-rich, extensin-like receptor kinase nucleic acids and polypeptides useful for increasing plant wounding or pathogen resistance.





Db 901 GCTGTGAAGCAGTTGAAAGTTGGGAGTGGTCAGGGAGAGAGGAGTTTCAGGCAGAGGTT 960  
Qy 321 GluIleIleSerArgValHiHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340  
Db 961 GAGATCATCAGCAGAGTTCCACACAGGCATCTGGTGTCTCTTGTGGTTATTGCATGCC 1020  
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAenAenAenLeuGluLeuHisLeu 360  
Db 1021 GGTGCCAAAGATTGCTTGTCTATGAGTTTGTCTTAAACAACAAATCTCGAGCTTCACCTC 1080  
Qy 361 HisGlyGluGlyArgProThrMetGluTyrPheSerThrArgLeuLysIleAlaLeuGlySer 380  
Db 1081 CATGCGGAGGACCGCCTACATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1140  
Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAenProLysIleIleHisArgAspIle 400  
Db 1141 GCTAAGGACTTCTTATCTTTCATGAAGATTGCAATCTTAAATCATTCACCGTGATATC 1200  
Qy 401 LysAlaSerAenIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420  
Db 1201 AAGGCTTCAAAACATATTATGATTTCAAGTTTGAAGCTAAGGTTGCTGTATTTGGTCTT 1260  
Qy 421 AlaLysIleAlaSerAspThrAenThrHisValSerThrArgValMetGlyThrPheGly 440  
Db 1261 GCTAAGATTGCTTCTGATACAAACACGCGATGATCAACACGCTGTGATGGGAACCTTTGGG 1320  
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460  
Db 1321 TACTTGGCTCGGAATAGCTGCAAGCGGAAGCTCACGGAGAGTCTGACGTTTCTCTCA 1380  
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAenVal 480  
Db 1381 TTTGGCGTGTGCTTTTGGAGCTCAATTACTTGGAGCTGACCCGTTGTATGCCAACAATGTC 1440  
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAenArgAlaSerGluGln 500  
Db 1441 TATGTAGATGACAGTTAGTTAGTGGCAGCACCATTGCTTAAACGAGCATCTGAGCAA 1500  
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAenAenGlyTyrAspArgGluMet 520  
Db 1501 GGAGACTTTGAGGGTTAGCTGATGCCAAAGATGAATAATGGGTATGCACAGAGAGAGATG 1560  
Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540  
Db 1561 GCTCGCTGGTGTGCTTGTGCTGCGGCTTGTGTGCGCCATTGAGCTGCGCCGAGACCTCGC 1620  
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAenValSerLeuSerAspLeuAenGluGly 560  
Db 1621 ATGACCCAGATTGTCGTCGTTAGAGGAAATGTATCACTGTCAAGATCTTAAACGAGGG 1680  
Qy 561 MetArgProGlyGlnSerAenValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580  
Db 1681 ATGAGACCAGGTCAAAGCAATGTATACAGCTCATACGGAGGAAGCACCGATTATGACTCG 1740  
Qy 581 SerGlnTyrAenGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600  
Db 1741 AGCCAGTACAATGAAGACATGAAGAAGTTTAGGAAAAATGGCACTTGGAACTCAAGAGTAC 1800  
Qy 601 AenAlaThrGlyGluTyrSerAenProThrSerAspTyrGlyLeuTyrProSerGlySer 620  
Db 1801 AACGCCACGGGTGAGTACAGTAATCCGACCACTGACTATGACTGTATGCCGCTGTTCA 1860  
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640  
Db 1861 AGCAGCGAGGGCAACACACCGCAATATGGAGATGGGGAAGATTAAAGAAACCGGTCTCAG 1920  
Qy 641 GlyTyrSerGlyProSerLeu 647  
Db 1921 GGTATATGTGGACCTTCTCTT 1941

RESULT 3  
ACF36549  
ID ACF36549 standard; cDNA; 2189 BP.

XX ACF36549;  
AC 18-DEC-2003 (first entry)  
XX B. napus PERK1 receptor kinase encoding cDNA.  
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;  
KW wound; pathogen resistance; plant growth; seed production; gene; ss.  
XX Brassica napus.  
XX Key Location/Qualifiers  
CDS 1..2189  
FT /\*tag= c  
FT /note= "ABR82938; this protein contains the amino acids  
FT corresponding to 5' and 3'UTR regions, though only the  
FT relevant aa residues (ABR82937) is used in the invention"  
FT 1..96  
FT /\*tag= a  
FT 97..2040  
FT /\*tag= b  
FT /product= "PERK1"  
FT /note= "ABR82937; proline-rich Extensin-like Receptor  
FT Kinase"  
FT 3'UTR 2041..2189  
FT /\*tag= d  
XX WO2003072763-A1.  
XX 04-SEP-2003.  
XX 28-FEB-2003; 2003WO-CA000274.  
XX 28-FEB-2002; 2002CA-02373903.  
XX 28-FEB-2002; 2002US-00086464.  
XX (GORI/) GORING D.  
PA (SILV/) SILVA N.  
PA (HAFF/) HAFFANI Y Z.  
XX Goring D, Silva N, Haffani YZ;  
XX WPI; 2003-712727/67.  
DR P-PSDB; ABR82937, ABR82938.  
XX Producing a transgenic plant having an increased plant resistance, plant  
PT growth or seed production comprises transforming a plant with a nucleic  
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase  
PT activity.  
XX Disclosure; Fig 1D; 123pp; English.  
XX The invention relates to producing a transgenic plant having increased  
CC plant height, number of branches, number of seed pods and/or seed  
CC production compared to a non-transgenic plant, and/or quicker flowering  
CC or later senescence compared to a non-transgenic plant. The method  
CC involves transforming a plant with a vector including a Proline-rich  
CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic  
CC acid molecule having PERK activity. The method, as well as the PERK  
CC nucleic acid molecule and polypeptide, are useful in increasing plant  
CC resistance to wounding and pathogens and in increasing plant growth and  
CC seed production. The nucleic acid molecule and polypeptide may also be  
CC used in producing transgenic plants or transgenic host cells. The present  
CC sequence represents a cDNA encoding a B. napus PERK1 receptor kinase  
CC polypeptide  
XX Sequence 2189 BP; 544 A; 592 C; 494 G; 559 T; 0 U; 0 Other;  
SQ

Alignment Scores:  
Pred. No.: 5e-93 Length: 2189  
Score: 3453.00 Matches: 647

Percent Similarity: 100.00%    Conservatives: 0  
Best Local Similarity: 100.00%    Mismatches: 0  
Query Match: 100.00%    Indels: 0  
DB: 9    Gaps: 0

US-10-086-464-2 (1-647) x ACF36549 (1-2189)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20  
Db 97 ATGTCCTCGGCGCGCTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAACTCCACA 156

Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProProSerSerProPro 40  
Db 157 ACCACCACTCTCTCCAGCTCCGCTCTCTCCACCACTCTCTCTCTCTCTCTCTCT 216

Qy 41 ProSerThrThrProProSerProProSerProProSerProProSerProProPro 60  
Db 217 CCATCCACTTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276

Qy 61 ProSerProProThrProSerThrProGlySerProProProProProProProPro 80  
Db 277 CCATCTCCACCACTCCATCTACGCGGGGATCTCCACCTCTCTCTCTCTCTCTCTCT 336

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrProPro 100  
Db 337 CCGCTCCCACTACGCGCGGATCTCCACCGCATCTGTACTCTCTCTCTCTCTCTCTCT 396

Qy 101 ProProSerValProGlyProProSerProSerProSerProSerProSerProPro 120  
Db 397 CCACCTTCAGTCCCGAGGACCGCTCAATCTCTCACCGGAGGAGGATCTCTCTGACCT 456

Qy 121 ProSerSerProSerProProSerProSerProSerProSerProSerProSerPro 140  
Db 457 CCATCT 516

Qy 141 GlyThrAlaThrGlyThrAlaLeuLeuValThrLeuLeuLeuLeuLeuLeuLeuLeu 160  
Db 517 GGAATCGCATCGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576

Qy 161 LysLysLysArgArgArgArgGluGluAspAlaThrValProProProProProPro 180  
Db 577 AAGAAGAACACGCGAGAGACGAAGAAGATCTTACTATGTCTCTCTCTCTCTCTCTCTCT 636

Qy 181 GlyProLysAlaGlyGlyProThrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGln 200  
Db 637 GGTCCCAAGCGGAGGACCTTACGCTGACGACGACGACGACGACGACGACGACGACG 696

Qy 201 ThrProProSerAspHisValThrSerLeuProProProProProProProProPro 220  
Db 697 ACACCCCGTCAGATCATGTCTGAGCTCACTACCCACCACTAAGGCTCCATCTCCA 756

Qy 221 ProArgGlnProProProProProProProProProProProProProProProPro 240  
Db 757 CCACGGCAACCTCTCTCCACCTCCACCCGCTTTCATGACGACGACGCGGCTCCGAC 816

Qy 241 TyrSerAspArgProValLeuProProProProProProProProProProProPro 260  
Db 817 TACTCGGACCGTCCAGTCT 876

Qy 261 SerThrPheThrThrGluLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 280  
Db 877 AGCACTTTCATACGAGGAGTACGTAGAGCCCAATGTTCTCCGAGGCGCAACTTG 936

Qy 281 LeuGlyGlnGlyPheGlyThrValHisLysGlyValLeuProProSerGlyLysGluVal 300  
Db 937 TTAGGACAAGCGCGGTTCGGTTCGTGCACAAAGGTGTGTTCCTAGTGGGAAAGAGTT 996

Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320  
Db 997 GCTGTGAAGCAGTTGAAGTTGGAGTGTCTCAGGAGAGAGGAGTTTCAGGCAAGGTT 1056

Qy 321 GluThrLeuSerArgValHisHisArgHisLeuValSerLeuValGlyThrCysIleAla 340

Db 1057 GAGATCATCAGCAGAGTTCCACACGAGCATCTGGTGTCTCTTGTGGTTATTGATCGCC 1116

Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuLeuHisLeu 360

Db 1117 GGTGCCAAAGATTTGCTTATGAGTTTCTTAAACAATCTCGAGCTTCACTC 1176

Qy 361 HisGlyGluGlyArgProThrMetGluThrSerThrArgLeuLysIleAlaLeuGlySer 380

Db 1177 CATCGCGAGGAGCGCTACAACTGAATGGAGCACCAGATTGAAGATTGCTTCTGGATCT 1236

Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400

Db 1237 GCTAAAGAGCTTTCTTATCTTCAAGATTGCAATCTCTAAATCAATTCACCGTATC 1296

Qy 401 LysAlaSerSerIleLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420

Db 1297 AAGCTTCAACAATATTGATAGATTCAAGTTTGAAGTAAAGTTGCTGATTTTGGTCTT 1356

Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440

Db 1357 GCTAAGATTGCTTCTGATACAAACACGATGTATCAACACGTTGATGGGAACCTTTGG 1416

Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460

Db 1417 TACTTGGCTCCGGAATACGCTCAAGCGGAAGCTCAGGAGAAGTCTGACGTTTCTCA 1476

Qy 461 PheGlyValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 480

Db 1477 TTTGGCGTTGCTTTTGGAGCTCATTAAGTGGAGCTGACCCGTTGATGCCAACAAATGC 1536

Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500

Db 1537 TATGTAGATGACAGCTTAGTTGACTGGCGACGACCATTTGCTTAACCGAGCATCTGAGCA 1596

Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520

Db 1597 GGAGACTTTGAGGTTTGTAGCTGATGCAAGATGAATAATGGGTATGACAGAGAGAGATG 1656

Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540

Db 1657 GCTCGCATGGTTGCTTGTGCTGGCGCTGTGTTCGCCATTCAGCTCGCGCAGACCTCGC 1716

Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560

Db 1717 ATGAGCCAGATTGTCGCTGCTTAGAAGGAAAATGATATCATCTGTCAGATCTTAACGAGGG 1776

Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580

Db 1777 ATGAGACCGGTCAAAGCAATGTATACAGCTCATACGAGGAGGAGCACCAGATTATGACTCG 1836

Qy 581 SerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600

Db 1837 AGCCAGTACAAATGAAGACATGAAGAAAGTTTAGGAAAATGGCACTTGAACCTCAAGAGTAC 1896

Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620

Db 1897 AACGCCACGGGTGAGTACAGTAATCCGACGAGTACTATGGACTGTACCCGCTGCTGTCA 1956

Qy 621 SerSerGluGlyGlnThrThrArgGluMetGlyLysIleLysArgThrGlyGln 640

Db 1957 AGCAGCGAGGCGCAACCAACACGCAATGGAGATGGGGAAGATTAAAGAGAACCGGTCA 2016

Qy 641 GlyTyrSerGlyProSerLeu 647

Db 2017 GGTTATAGTGGACCTTCTCTT 2037

RESULT 4  
ACF36557  
ID ACF36557 standard; DNA; 1959 BP.  
XX ACF36557;  
AC ACF36557;  
XX ACF36557;  
DT 18-DEC-2003 (first entry)

XX	A. thaliana PERK1 protein encoding genomic DNA.
DE	
XX	PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW	wound; pathogenesis resistance; plant growth; seed production; gene; ds.
XX	
XX	Arabidopsis thaliana.
OS	
XX	WO2003072763-A1.
FN	
XX	04-SEP-2003.
PD	
XX	
PF	28-FEB-2003; 2003WO-CA000274.
XX	
PR	28-FEB-2002; 2002CA-02373903.
XX	
PR	28-FEB-2002; 2002US-00085464.
XX	
PA	(GORI/) GORING D.
XX	(SILV/) SILVA N.
PA	(HAFF/) HAFFANI Y Z.
XX	
PI	Goring D, Silva N, Haffani YZ;
XX	
DR	WPI; 2003-712727/67.
XX	
PT	Producing a transgenic plant having an increased plant resistance, plant
FT	growth or seed production comprises transforming a plant with a nucleic
PT	acid molecule having a Proline-rich Extensin-like Receptor Kinase
PT	activity.
XX	
PS	Disclosure; Fig 21d; 123pp; English.
XX	
CC	The invention relates to producing a transgenic plant having increased
CC	plant height, number of branches, number of seed pods and/or seed
CC	production compared to a non-transgenic plant, and/or quicker flowering
CC	or later senescence compared to a non-transgenic plant. The method
CC	involves transforming a plant with a vector including a Proline-rich
CC	Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
CC	acid molecule having PERK activity. The method, as well as the PERK
CC	nucleic acid molecule and polypeptide, are useful in increasing plant
CC	resistance to wounding and pathogens and in increasing plant growth and
CC	seed production. The nucleic acid molecule and polypeptide may also be
CC	used in producing transgenic plants or transgenic host cells. The present
CC	sequence represents an A. thaliana PERK1 DNA (At3g24550)
XX	
SQ	Sequence 1959 BP; 496 A; 514 C; 425 G; 524 T; 0 U; 0 Other;
 Alignment Scores:	
Pred. No.:	5,28e-78 Length: 1959
Score:	2936.50 Matches: 574
Percent Similarity:	90.61% Conservative: 24
Best Local Similarity:	86.97% Mismatches: 41
Query Match:	Indels: 21
DB:	Gaps: 12
 US-10-086-464-2 (1-647) x ACF36557 (1-1959)	
Qy	1 MetSerSerAlaProSerProGlyThr-----GlySerProSerProSerProSer 17
Db	1 ATGTCCACGCGCGTCTCCAGGCCTACTCCATCACCATTCTCCACGGTCTCTGCCACA 60
Qy	18 AsnSerThrThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db	61 AACTCGACACACCACCACTCTCTCTCAGCAGCCTCTTCTCTCTCCACACCACTCT 120
Qy	36 SerSerProProPro---ProSerThrThrPro-----ProSerProProSerSerArgSer 54
Db	121 TCCTTCCTCTCGTCGCGGTCTCACTTAATCAACTCTCTCTCTCTCTCTCTTTA 160
Qy	55 ThrProSerAlaProProProSerProProThrProThrProThrProGlySerProPro 74
Db	181 CCTCCCTCTCTCTCTCACTCTCTCTCTCGATCTTTAACT-----CCTCCT 228

```
Db 1297 AACACTCATGTATCTACACGGCTGATGGAACTTTGGGTATTTGGCTCCGGAATATGCT 1356
Qy 448 AlaserGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1357 GCAAGTGGAAAGCTCACAAGAAAGTCTGACGTTTTCTCATTTGGCGTTGTACTTTTGGAA 1416
Qy 468 LeuIleThrGlyArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db 1417 CTTATTACTGGAGGCGCCTGTTGATCGGAACATGCTATGTAGATGACAGCTTAGTT 1476
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1477 GACTGGGCACGACCATTCCTTAACCGAGCATCTCGAGGAAGGAGATTTTCAGGGTTTGCT 1536
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1537 GATTCAAGATGGGTAATGATGATGACAGAGAGAGATGGCTCGCATGTTGCTTGGCGCT 1596
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1597 GCGGCTTGTTGGCATTGAGCTCGCCGAGACCTCGCATGAGCCAGATAGTACGGCG 1656
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1657 TTAGAAGGAATGTATCGCTGCTGATCTTAACGAAGGAGATGAGACCGGGTCACAGCAAC 1716
Qy 568 ValTyrSerSerTyrGlyLysThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1717 GTATACAGCTCATATGGAGGAAGCACACAGACTATGACACGCGCAATACACAGCAGCATG 1776
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1777 ATAAAGTTAGAAAATGGCTCTTGAACTCAAGNATACGCACACCGCGGAGTACAGT 1836
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
Db 1837 AATCCAACAGTACTACGAGCTGTACCCGCTCTGTTCAAGCAGTGAAGGTCAAGCCACA 1896
Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlnGlyTyrSerGlyProSerLeu 647
Db 1897 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGGTCAAGGTATATAGTGGACCTCTCTT 1956

RESULT 5
ID AAC36968 standard; DNA; 2324 BP.
XX AC AAC36968;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15710.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway; metabolic pathway;
XX KW promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX FN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 29-MAR-1999; 99US-0126284P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
```

Page 8







Db	1276	GGCTACAGTGGTGCTACAGCTCAGGATACAGCGAGCCTCG	1317	Db	63	GATGGCTTCTCTGATGCCAACCTCTTGGACAAGAGGATTTGGATATGTGCACAGAGA	122
RESULT 8							
ACF36556							
ID	ACF36556	standard; DNA; 1488 BP.		Qy	293	ValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGly	312
XX	XX	XX		Db	123	ATTCTTCCCAACGCGAAGGAGTGGCAGTGAAGCAATTGAAGGCTGGAAGCGGCGAAGGG	182
AC	ACF36556;			Qy	313	GluArgGluPheGlnAlaGluValGluLeuLeuLeuSerArgValHisArgHisLeuVal	332
DT	18-DEC-2003	(first entry)		Db	183	GAGCGTGAATTTCCAAAGCTGAAGTTGAGATTAATTAAGCCGTGTCATCAACAACATCTTGT	242
DE	G. max	PERK partial EST sequence.		Qy	333	SerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValPro	352
XX	PERK1;	transgenic; plant; proline-rich extensin-like receptor kinase;		Db	243	TCITTTGGTTGATACTGCATCAGTGGTCCAGAGGCTGCTTGTATTGAATTTGTTCC	302
KW	wound;	pathogen resistance; plant growth; seed production; soybean; gene;		Qy	353	AsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTyrSerThr	372
KW	ds.			Db	303	AACAACACATTTGGAATTTCCATTTGCATGGAAAAGGAGACCTACCATGTGGCCACACA	362
OS	Glycine max.			Qy	373	ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsn	392
XX	WO2003072763-A1.			Db	363	AGACTAAGAATTTGCTTTAGGATCTGCTAAGGAGCTGGCGTATCTTCATGAAGATTGTCAT	422
PN	04-SEP-2003.			Qy	393	ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu	412
XX	28-FEB-2003; 2003WO-CA000274.			Db	423	CCTAAGATCATCCATCGTATATCAAGCTGCCAACATCTCTCTGGATTTTAAGTTGAA	482
PR	28-FEB-2002; 2002CA-02373903.			Qy	413	AlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer	432
PA	(SILV/) SILVA N.			Db	483	GCAAAAGGTTGCAGATTTCCGCTTTCGAAAGTTTCTTCTGATGTCAATACTCATGTTTCT	542
PA	(HAFF/) HAFFANI Y Z.			Qy	433	ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu	452
XX	Goring D, Silva N, Haffani YZ;			Db	543	ACTCGAGTGATGGGAGCTTTTGGGTATTTGGCTCCAGAATAATGCTTCTAGTGGAAAAC	602
DR	WPI; 2003-712727/67.			Qy	453	ThrGlyLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg	472
XX	Producing a transgenic plant having an increased plant resistance, plant			Db	603	ACACACAATTCAGATGTTTCTCTACGGAGTCATGCTCTCCAGATTAAATAACCGGACGA	662
PT	growth or seed production comprises transforming a plant with a nucleic			Qy	473	ArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTyrPheAlaArgPro	492
PT	acid molecule having a Proline-rich Extensin-like Receptor Kinase			Db	663	CGGCGCGTCGATAAAATCAAACTTTCATGAGGATAGTTTGGTAGACTGGCGTAGGCCT	722
XX	activity.			Qy	493	LeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAspAlaLysMetAsn	512
PS	Disclosure; Fig 21c; 123pp; English.			Db	723	TTGCTCACACAGCTTTGGAAAGAGGATTTTGAATTCATATTATGACCCAAAGGCTCCAG	782
CC	The invention relates to producing a transgenic plant having increased			Qy	513	AsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg	532
CC	plant height, number of branches, number of seed pods and/or seed			Db	783	AATGACTATGATCCTAATGAGATGGCAGCAATGGTGGCTTCTGCTGGCGCTTGCAATCGT	842
CC	production compared to a non-transgenic plant, and/or quicker flowering			Qy	533	HisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal	552
CC	or later senescence compared to a non-transgenic plant. The method			Db	843	CATTTCGCAAAAGCGTCGACCAAGGATGAGCAGGTTGTCCGCGCTCTCGAAGGAGATGC	902
CC	involves transforming a plant with a vector including a Proline-rich			Qy	553	SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr	572
CC	Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic			Db	903	TCTCTAGCAGATCTGAATGAAGGAATAAGACCTGGACACACACTATGTATAGTTCTCAT	962
CC	acid molecule having PERK activity. The method, as well as the PERK			Qy	573	GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys	592
CC	nucleic acid molecule and polypeptide, are useful in increasing plant			Db	963	---GAAAGCTCAGATTATGACATCGACAGTACAAAGGAGAGACATGAAAAGTTCCAGGAAA	1019
CC	resistance to wounding and pathogens and in increasing plant growth and			Qy	593	MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp	612
CC	seed production. The nucleic acid molecule and polypeptide may also be			Db	1020	ATGGCATTGGGAACTCAGGAGTATGCTGTCGAACAGTACAGTACGTCGCGCTACAAAGTGAG	1079
CC	used in producing transgenic plants or transgenic host cells. The present			Qy	613	TyrGlyLeuTyrProSerGlySerSerGluGly-----GlnThrArgGlu	629
CC	sequence represents a G. max PERK EST (expressed sequence tag) partial			Db	1080	TATGGTTTAAACCATCAGGCTCAAGTAGTGAAGCAGACAGCCGCCAACCCACCAAGGGA	1139
CC	sequence (TIGR Accession No. TC123698)			Qy	630	MetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSer	646
XX							
SQ	Sequence 1488 BP; 412 A; 291 C; 348 G; 437 T; 0 U; 0 Other;						
Alignment Scores:							
Pred. No.:	3,246-40	Length:	1488				
Score:	1637.50	Matches:	312				
Percent Similarity:	89.67%	Conservative:	44				
Best Local Similarity:	78.59%	Mismatches:	36				
Query Match:	47.42%	Indels:	5				
DB:	9	Gaps:	3				
US-10-086-464-2 (1-647) x ACF36556 (1-1488)							
Qy	253	GlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThr	272				
Db	3	GGAATTTCAITGGGGTTCTCTAAGAGACATTCATGATAGAGGATTTGGCAGCGCAACT	62				
Qy	273	AsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGly	292				

Db 1140 ATGGAATGAGAAATGAAG---AACAAATCAAGGTTTCAGTGGAAAGTTCT 1187  
RESULT 9  
AAF77096  
ID AAF77096 standard; DNA; 2520 BP.  
XX  
XX AC  
XX AAF77096;  
DT 17-MAY-2001 (first entry)  
XX  
XX Arabidopsis gene #2.  
DE Arabidopsis  
XX  
XX KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.  
XX OS Arabidopsis thaliana.  
XX  
XX PN W0200114563-Al.  
XX  
XX PD 01-MAR-2001.  
XX  
XX PF 18-AUG-2000; 2000WO-CA000966.  
XX  
XX PR 19-AUG-1999; 99US-0149466P.  
XX PR 13-OCT-1999; 99US-0159122P.  
XX  
XX (GORI/) GORING D.  
PA (SILV/) SILVA N.  
XX  
XX Goring D, Silva N;  
XX  
XX WPI; 2001-244305/25.  
XX  
XX PT New proline-rich, extensin-like receptor kinase nucleic acids and  
PT polypeptides useful for increasing plant wounding or pathogen resistance,  
PT or for producing transgenic plants with increased wounding or pathogen  
PT resistance.  
XX  
XX Example; Fig 12; 91pp; English.  
XX  
XX PS The present invention relates to proline-rich extensin-like receptor  
CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for  
CC increasing the resistance of plants to wounding and pathogens. These are  
CC also useful for producing transgenic plants with increased wounding and  
CC pathogen resistance compared with a wild type plant, as well as in assays  
CC for identifying and developing compounds to inhibit and/or enhance  
CC polypeptide function directly  
XX  
XX SQ Sequence 2520 BP; 717 A; 553 C; 537 G; 713 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 2.93e-36 Length: 2520  
Score: 1507.00 Matches: 368  
Percent Similarity: 53.47% Conservative: 79  
Best Local Similarity: 44.02% Mismatches: 146  
Query Match: 43.64% Indels: 244  
DB: 4 Gaps: 22  
  
US-10-086-464-2 (1-647) x AAF77096 (1-2520)  
Qy 3 SerAlaProSerProGlyThrGlySerProProSerProSerProSerProSerThrThr 22  
Db 36 TCTGCTCTCCAAACAATCCACCTCTCTCCATCTCCACCGCTAATACCAATTCACAC 95  
Qy 23 ThrProProAlaSerAlaProProProThrThrProSerProSerProProSer 42  
Db 96 ACCTCTCTCCG-----CCGGCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAAGGA 146  
Qy 43 ThrIleProThrProProSerSerArgSerThrProSerAlaProProSer 62  
Db 147 GACTCATCATCATCGCCACCTCTCTGATTCACATCTCCACCGCTCCACAAAGCTCTTAAC 206  
Qy 63 ProProThrProSerThrProGlySerProProProLeuProGlnProSerProAla 82

Db 207 CTTCTTAAT---TCCTCTAATAAATCTCTCTCCCTCCGTACAGGGCGGTGGAGGAA 263  
Qy 83 ProThrThrProGlySerProProAlaProValThrProProThrArgAsnProProPro 102  
Db 264 AGAGGAAATGGAGAAACAATGGTGGCAATGATATCCACCGTCACCGGGCTCTCTCTCT 323  
Qy 103 SerValProGlyProProSerAsnProSerArgGlyGlySerProArg-----Pro 120  
Db 324 TCT-----CCTCTCTTAGGAGTAATGGAGATAATGGTGTAGCAGATCATCGCCA 374  
Qy 121 ProSer-----SerProSerProProSerPro-----SerSerAsp 132  
Db 375 CCAGGAGACACTGGAGGCTCTCGTCTAGACAACTCTCTCTAGCGGAGGAGCAGTGA 434  
Qy 133 Gly-----LeuSerThrGlyValValGlyValValGlyValIleAlaIleGly 145  
Db 435 GGAGGTGGAGGTGAAGAAGTAATACGATACGCGATCATAGTTGGTGTATTAGTCGA 494  
Qy 146 GlyValAlaLeuLeuValIleValThrLeuLeuCysLeuLeuCysLysLysLysArgArg 165  
Db 495 GCTGACTTTTGTATGATCGTTCTTATTATTGTGTGTCTTAGACGCAAAAGAGAGAAA 554  
Qy 166 ArgAspGluGluAspAlaTyrTyrValProProProProProGlyProGlyAla--- 184  
Db 555 -----GACTCTCTCTAC-----CCTGAACCCATGAAGGTAAACATATAC 596  
Qy 185 -----GlyGlyPro 187  
Db 597 ACACCTTATGTTTCAACAAATAGAAAGCTTAGATTCTTTCAATAAAATTTTCAGAAACCA 656  
Qy 188 TyrGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisVal 207  
Db 657 ATATCAATACTATGGAACAACAACAACAATGCTTACAGAAATTTATCCGAA----- 710  
Qy 208 ValThrSerLeuProProProProProProProProProPro----- 221  
Db 711 -----TTGGCACCTAAATTCACAAAGGCCAAACCAATCTACTGTGTGTGGG 761  
Qy 222 ArgGln-ProProProProProProPheMetSerSerSerGly-----GlySe 239  
Db 762 AGCGGTGGACATCATCCGCTCTCTCTCGCGATCCCTACAGCGGAGAGATTCTTTC 821  
Qy 239 rAspTyrSer-----AspArgProValLeuProProProProGlyLeuValLeu 257  
Db 822 CATGTACTCAGGCCCATCATCGCCAGTTTACCTCTCTCTCTCTCTCTCTCTCTCG 881  
Qy 257 yPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerG 277  
Db 882 ATTCAACAGAGCACATTTTACTTACCAAGAGCTTGGCGCTGCAACAGAGGGTTTACGGA 941  
Qy 277 uAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisGlyValLeuProSerG 297  
Db 942 TGCTAACCTTTGGGACAGGAGGATTTGGGTATGTCCATAAAGAGTCTTTCCTTAGCGG 1001  
Qy 297 YLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheG 317  
Db 1002 GAAAGAGTAGCAGTTAAGAGTTTAAAGCGGTAGCGGACAGAGAGGAGTTTCA 1061  
Qy 317 nAlaGluValGluIleLeuSerArgValHisArgHisLeuValSerLeuValGlyTyr 337  
Db 1062 AGCTGAGTCCATATCATATTAGCCGTGTGCATCATCGGTATCTTCTTTCTTTGGTGGATA 1121  
Qy 337 rCysIleAlaGlyAlaLysArgLeuLeuValThrGluPheValProAsnAsnLeuG 357  
Db 1122 TTGCATAGCTGATGACAGAGGATTTGGTTTATGAGTTTGTCTTCTTAACTTTTGA 1181  
Qy 357 uLeuHisLeuHisGly----- 362  
Db 1182 ATATCATCTTCATGGTTAGACCACCTTAAACCTTTGAGTACTAAGTTTATTTCTTAAT 1241  
Qy 363 -----Glu-GlyArgProThrM 368  
Db 1242 CTATATATCAAGAAAGTTGTGAACCTTTAAATTTGTTGTTAGGGAATAAATCTTCGCTAA 1301

```
QY 368 etGluTrpSerThrArgLeuIysIleAlaLeuGlySerAlaIysGlyLeuSerTyrLeuH 388
Db 1302 TGGAGTTCTCCACTAGTTGGGTATCGCTTAGGTGCTCGAAAGGACTCGTTACCTTC 1361
QY 368 leGluAspCys----- 391
Db 1362 ACGAAGACTGTAAAGTTTTAAACATTCCACCATCTCATTTTCTTAACCAAGTTGCATAAAAC 1421
QY 392 -----Asn 392
Db 1422 AGAGAAAGCTCTGCTCTGACTAGTGTATTCTTTTGGCTGAGAAAATGGTGAGGCCAT 1481
QY 393 ProLysIleIleHisArgAspIleLysAlaSerAsenIleLeuIleAspPheLysPheGlu 412
Db 1482 CCTCGGATCATTCACCGGACATCAAGCTCGCAATATATCTCTTGGGACTTCAACTTTGAT 1541
QY 413 Ala----- 413
Db 1542 GCTATGGTGATAAACTAGTAGTTCGCATTCTACCTAGCGTTTGTGTTAAGACTACATTGA 1601
QY 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAs 426
Db 1602 TGACATTTTGCATTTGTTATTTCAGGTGGCTGATTTTGGATTAGCTAAGTTAAACATCTGA 1661
QY 426 pThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db 1662 TAACAACACTCATGTATCTACTCGTGTGATGGGAACCTTTCGGGTAAAGCGTTTACCCTAT 1721
QY 440 ----- 440
Db 1722 GATAAGATTGTCGTGACACTCAAGAAACATAACCTTTGTAGACTAATCTACTTGTCTTC 1781
QY 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrG 454
Db 1782 TTTCCACAAACATGTGTAGATATCTAGCTCCAGAAATATGCTTCAAGCGGTAAATTAACCG 1841
QY 454 luLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgP 474
Db 1842 AGAAATCCGATGTTTTCTCTTACGGAGTATGTGTTATTCGAACTTATACTGGAAACGAC 1901
QY 474 roValAspAlaAsnValTyrValAspArgGluMetAlaArgMetValAlaCysAlaAla 489
Db 1902 CGGTTGAT---AATAGCATCACCATTGGACGACACCTTAGTAGATTGGGTATTCATGCATG 1958
QY 489 ----- 489
Db 1959 TAACATATGTATCGTGTATATATGTTTTCGGCTTTTTCGGCTACTAATGATCATGAATA 2018
QY 490 ---AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508
Db 2019 CAGGCTCGGCTCTTATCGCTCGCGCTAGAGATGGAACCTTTAATGAGTCGCGAGAT 2078
QY 509 AlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAla 528
Db 2079 GCGAGCTTGAAGGCACTACAAACCCGCAAGAAATGGCTCGAATGGTGACTTGTGCGCT 2138
QY 529 AlaCysValArgHisSerAlaArgArgProArgMetSer----- 542
Db 2139 GCTAGCATTCGTCATTTCGGGGGTAAACGTCCTCAAGATGAG-CCAGGTGAATCAAAATTA 2197
QY 543 -----GlnIleVal 545
Db 2198 TAACATAAAGTCTATTTTTTGTGAGAGAATAACAAACAAATGTGTGTTTTTTCAGATAGTA 2257
QY 546 ArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGln 565
Db 2258 AGAGCGTTAGAGGAGAGATGCTTCTTAGATGCTTTAAACGAGAGGTGTGAAGCCAGGACAC 2317
QY 566 SerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGlu 585
Db 2318 AGTAACGTTTACGGGTCAATTGGGCAAGCTCGGATTATAGTCAGACATCTTACAATGCA 2377
```

```
QY 586 AspMetLysLysPheArgLysMetAlaIeuGlyThrGlnGluTyr-----AsnAla 602
Db 2378 GACATGAAGAAATTCAGACAGATAGCTTTGTCGAGCCAAAGAAATTCACAGTCAGTGACTGT 2437
QY 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGly 614
Db 2438 GAAGGAACATCTAGTAATGATTCTTAGAGATATGGGA 2473

RESULT 10
ACF36551
ID ACF36551 standard; DNA; 2520 BP.
XX
AC ACF36551;
XX
DT 18-DEC-2003 (first entry)
XX
DE Arabidopsis PERK1 receptor related protein encoding DNA.
XX
KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
KW wound; pathogen resistance; plant growth; seed production; gene; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO2003072763-A1.
XX
PD 04-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-CA000274.
XX
PR 28-FEB-2002; 2002CA-02373903.
XX
PR 28-FEB-2002; 2002US-00086464.
XX
PA (GORI/) GORING D.
PA (SILV/) SILVA N.
PA (HAFF/) HAFFANI Y Z.
XX
PI Goring D, Silva N, Haffani YZ;
XX
DR WPI; 2003-712727/67.
DR P-PSDB; ABR82940.
XX
PT Producing a transgenic plant having an increased plant resistance, plant
PT growth or seed production comprises transforming a plant with a nucleic
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
PT activity.
XX
PS Disclosure; Fig 12; 123pp; English.
XX
CC The invention relates to producing a transgenic plant having increased
CC plant height, number of branches, number of seed pods and/or seed
CC production compared to a non-transgenic plant, and/or quicker flowering
CC or later senescence compared to a non-transgenic plant. The method
CC involves transforming a plant with a vector including a proline-rich
CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
CC acid molecule having PERK activity. The method, as well as the PERK
CC nucleic acid molecule and polypeptide, are useful in increasing plant
CC resistance to wounding and pathogens and in increasing plant growth and
CC seed production. The nucleic acid molecule and polypeptide may also be
CC used in producing transgenic plants or transgenic host cells. The present
CC sequence represents a PERK1 polypeptide related protein encoding genomic
CC DNA from A. thaliana (Accession NO. AAD15491)
XX
SQ Sequence 2520 BP; 717 A; 552 C; 538 G; 713 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.58e-36 Length: 2520
Score: 1504.00 Matches: 367
Percent Similarity: 53.47% Conservative: 80
Best Local Similarity: 43.90% Mismatches: 146
Query Match: 43.56% Indels: 22
DB: 9 Gaps: 22

US-10-086-464-2 (1-647) x ACF36551 (1-2520)
```



```

Db 2139 GCTAGCATTCGTCATTCGGGGCTAAACGTCCAAAGATGAG-CCAGGTGAATCAAAATTA 2197
Qy 543 -----GlnIleVal 545
Db 2198 TAACTAAAGTCTATTTTGTTCAGAGATAACAAACAATGTTGTGGTTTTCAGATAGTA 2257
Qy 546 ArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGln 565
Db 2258 AGAGCGTTAGAGGAGAAAGTGTCTTAGATGCTTTAAACGAAGGTGTGAAGCCAGGACAC 2317
Qy 566 SerAsnValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGlu 585
Db 2318 AGTACGTTTACGGGTCAATGGGACGACGCTCGGATTATAGTCAGACATCTTACAATGCA 2377
Qy 586 AspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyr-----AsnAla 602
Db 2378 GACATGAGAATAATTCAGACAGATAGCTTTGTGAGCCAGAAATCCCAAGTCAGTGACTGT 2437
Qy 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGly 614
Db 2438 GAAGGAACATCTAGTAATGATTTCTAGAGATATGGGA 2473

```

## RESULT 11

AAF77097 standard; DNA; 2820 BP.

```

AC AAF77097;
DT 17-MAY-2001 (first entry)
XX Arabidopsis gene #3.
DE Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
XX Arabidopsis thaliana.
XX WO200114563-A1.
XX 01-MAR-2001.
XX 18-AUG-2000; 2000WO-CA000966.
XX 19-AUG-1999; 99US-0149466P.
XX 13-OCT-1999; 99US-0159122P.
XX (GORI/) GORING D.
XX (SILV/) SILVA N.
XX Goring D, Silva N;
XX WPI; 2001-244305/25.
XX New proline-rich, extensin-like receptor kinase nucleic acids and
PT polypeptides useful for increasing plant wounding or pathogen resistance,
PT or for producing transgenic plants with increased wounding or pathogen
PT resistance.
XX Example; Fig 13; 91pp; English.

```

CC The present invention relates to proline-rich extensin-like receptor  
 CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for  
 CC increasing the resistance of plants to wounding and pathogens. These are  
 CC also useful for producing transgenic plants with increased wounding and  
 CC pathogen resistance compared with a wild type plant, as well as in assays  
 CC for identifying and developing compounds to inhibit and/or enhance  
 CC polypeptide function directly

SQ Sequence 2820 BP; 819 A; 639 C; 577 G; 795 T; 0 U; 0 Other;

Alignment Scores: 2e-35 Length: 2820  
 Pred. No.: 1479.50 Matches: 367  
 Score:

```

Percent Similarity: 50.00% Conservative: 99
Best Local Similarity: 39.38% Mismatches: 151
Query Match: 42.85% Indels: 318
DB: 4 Gaps: 28

US-10-086-464-2 (1-647) x AAF77097 (1-2820)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProSerProSerProSerAsnSerThr 20
Db 52 GTGGAATTCATCTCCTGCCCTGAAACC---TCAATGGGACACCAACCGTCAACCGGAACA 108
Qy 21 ThrThrThrProProAlaSerAlaProProThrThrThrProSerSerProProPro 40
Db 109 TCG-----CCGTCTAATGAGTCATCGCGGCCCAACACCACTCTTTCACCAACCA 159
Qy 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro 60
Db 160 TCATCA-----ATATCTGCTCCTCCGCCAGATATCTCCGCTCTCTTTTACCGCG 210
Qy 61 ProSerProProThr-----ProSerThrProGlySerProProProPro 75
Db 211 CCTGCACCAACGCGCAAGAAACGTCACCTCTCATATCTCCTGCTCTCATCGCGCTGTT 270
Qy 76 -----ProGln-----ProSerProProAlaPro-----Thr 84
Db 271 GTAGCTAATCCGTCACCGCAGACTCCAGAGAAATCTTCTCCACCTGCACCTGAAGGCTCA 330
Qy 85 ThrProGlySerProProAla---ProValThrPro-----ProThrArgAsn 99
Db 331 ACTCCTGTAAAGCCACCTGCACCAACCAACACCGTCGACCAATCACCAGGAAGACCA 390
Qy 100 ProProPro-----Pro-----Pro-----Pro----- 102
Db 391 ACTCCTCTCTCTCTGTCGCAATGATGACCGAAACAGAAACCAATGGCGGAAACAAACAAC 450
Qy 103 -----SerValProGlyProProSerAsnProSerArg-----GluGlyGly 116
Db 451 AGAGACGGCTCCACCACTACCAACCGTCGTCGAGGAACAGAACTTCGGGTGACGGTGGC 510
Qy 117 SerProArgProProSerSerProProSerProSerSerAspGlyLeuSerThr 136
Db 511 TCACCTTCACCACTCGTGTGATAAGCCCTCTCAGAAATAGTGGAGATTCAGACTCATCA 570
Qy 137 -----GlyValValValGlyIleAlaIleGlyGlyVal 147
Db 571 TCGGGTAATCATCAAGCAACATTCGATGATTGATTTGGAGTCTTGTAGGAGACGGG 630
Qy 148 AlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArgArgAsp 167
Db 631 CTTTGTCTTCTACTTGACAGTGTATTGTCATCTGTTGCAACAGAGAGAAAGAAAGAA 690
Qy 168 GluGluAspAla-----TyrTyrValProProProProProGlyProLysAla 184
Db 691 TCTCCTCAGGTCAACCACTGCTACTTACAATAACAATCCTTATGGAGGAGACCCCTCA 750
Qy 185 Gly----- 185
Db 751 GGTAAATTACAGTTTAGTAACTCGAATTTAATTTGTAGCCTTAATGGTGTGTTTAGGT 810
Qy 186 -----GlyProTyrGlyGlyGlnGlnGlnGln 194
Db 811 TTCAGACGATCATAGTCTAATGGTTTCTGCTAGCTCCATAATGGCAAAAGGATTAGATT- 869
Qy 195 TrpArgGlnGlnAsnAlaThr-----ProPro 203
Db 870 TATAAGCTAAAGAGAGATGTTTCATAGTAGTAGTAAATGCTGTTATTATCAAGGGAACACCT 929
Qy 204 SerAspHisValValThrSerLeuProProProProLysAlaProSerProProArgGln 223
Db 930 CAAGATCATGTGGTG---AATATGGCTGGTCAAGAGGTGGGAAATGGGGTCCACAGCAA 986
Qy 224 ProProProProProProProPheMetSerSerSerGlyGlySerAspTyrSerAsp 243

```

Db 987 CCTGTGCTGTGCTCT- - - - - CACAGTGATGCTTCCAACTTAACCGGT 1028  
Qy 244 ArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysThrPhe 263  
Db 1029 CGAACTGTAACCTGACCTCAA- - - - - CAGTCACTCTTGGTCAACAAGCACTTTC 1085  
Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuGlyGln 283  
Db 1086 ACATACGATGAATGCTCATTGCAACAGAGGTTTCGCTCAGTCAAAATTTGCTAGGACAA 1145  
Qy 284 GlyGlyPheGlyTyrValHisGlyValLeuProSerGlyLysGluValAlaValLys 303  
Db 1146 GGAGATTGGGTATGTTTCATAAAGGAGTCTCGCTAGTGGCAAGAAGTTGCAAGTGAAG 1205  
Qy 304 GlnLeuLysValGlySerGlyGlnGluArgGluPheGlnAlaGluValGluLeuLe 323  
Db 1206 AGTCTTAAACTGGAGTGGACAGGGGAAAGCGAGTTTCAAGCAGAGTTGATATCAT 1265  
Qy 324 SerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343  
Db 1266 AGCGGTGTCATCATCGTCACTCGTCTCTGTTGGATATGTCATCTCTGTTGGTCAA 1325  
Qy 344 ArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHis 361  
Db 1326 AGACTTTTGGTTATGAGTTTATACCTAACCAACACTCTTTGAATTTCAATCATGATCA 1385  
Qy 361 - - - - - 361  
Db 1386 TTCATCTAACAGATGTTTCTGTATTAACAAAACCTTTAAGTATGTTTCTCTTTAAT 1445  
Qy 362 - - - - - GlyGluGlyArgProThrMetGluTrpSerThrArgLe 374  
Db 1446 CAGGAACATGATTGAAATTTTCAGGAAGAGGTCGTCGGTCTTGGATTGGCTCACAAGAGT 1505  
Qy 374 uLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluLeuPhe 391  
Db 1506 GAAGATTGCAATGGGATCAGCTAGAGGCTTTGCATATTTGTCATGAAGACTGTAAAGAAAT 1565  
Qy 391 - - - - - 391  
Db 1566 CTTTATCTCACATATTGTCATCAGTTTCTATCTGCTCTCTACAAATATTGAAAGATTGT 1625  
Qy 392 - - - - - AsnProLysIleIleHisArgAspIleLysAlaSerAs 404  
Db 1626 ATATGTTACATCAATATATAGTCAACCTCGCATTTATCCACAGATATCAAGCTGCAAA 1685  
Qy 404 nIleLeuLeuAspPheLysPheGluAla- - - - - 413  
Db 1686 CATTCCTTCTGATTTTCAGTTTGTAGAC- CAAGGTATGTGTGTATATATATCGACTCTTGTATC 1744  
Qy 413 - - - - - 413  
Db 1745 TACTTTTACTTTCATTTGTTCTCATTTTGTGTTTCCAAATCTGTGTCGATGTTGATCAGT 1804  
Qy 414 - - - - - LysValAlaAspPheGlyLeuAlaLysIleAlaSerAspTh 427  
Db 1805 CTTATTGTTGTAATATATGTCAGTGGCAGATTTTGGATTGGCTAAGCTATCTCAAGACAA 1864  
Qy 427 rAsnThrHisValSerThrArgValMetGlyThrPhe- Gly- - - - - 440  
Db 1865 CTATACTCATGTCCTCCACTCGCTCATGGGAACCTTTTGGGTAAAGCAGCTTTGTTAAATGT 1924  
Qy 440 - - - - - 440  
Db 1925 CTCAACTCATCCACATTAATTAGTTTCTTTCATCTGTTTTCATTTTAAACATTTCTTGGATTCA 1984  
Qy 441 --TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheS 460  
Db 1985 GATATTAGTCTCCAGAGTATGATCAACAGCGGAAGTATTCGACAAATCTGATGTTTCT 2044  
Qy 460 erPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480  
Db 2045 CATTTGGAGTAATGCTTCTTGAGCTCATAAACCGGAAGACCTCTCTCGATCTAACTCGA- 2103

Qy 480 alTyrValAspAspSerLeuValAspTyr- - - - - 489  
Db 2104 --GAATGGAGATAGCTTGGTAGATTGGGTAAAGTCGGTCCCGCTCTTCGGTTTACTT 2161  
Qy 489 - - - - - 489  
Db 2162 GTTTTAATCCCAAAACACTTTTCCAAAGCAAAAACAGAAACAAATCTTACTATTGTTGTTC 2221  
Qy 490 --AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspA 509  
Db 2222 AGCAAGCGCTTGTGTTTGAAGCAGCTCAAGATGGAGATTACAACCAATTTGGCTGATC 2281  
Qy 509 laLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaA 529  
Db 2282 CACGCTAGAGCTAAACTACAGTCATCAAGAGATGGTTCAATGGCTTCTTGTGCGAGCTG 2341  
Qy 529 laCysValArgHisSerAlaArgArgProArgMetSerGlnIle- - - - - 544  
Db 2342 CAGCAATCAGACATTCAGCAAGAGAGCGGCTTAAGATGAGCCAGGT-TCAAAAAACTCATA 2400  
Qy 544 - - - - - 544  
Db 2401 CCACCTGTGTTGTTCTATGTTATATTTTACTCAAAATTAATCTTGATGATAAATGTGAC 2460  
Qy 545 - - - - - ValArgAlaLeuGluG 550  
Db 2461 ATACTAATGAATCTTGAACATGTGTATGGTAAATGAAAGATTGTACGAGCACTAGAG 2520  
Qy 550 lyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrS 570  
Db 2521 GAGATATGTCAATGGATGATCTAAGTGGGGAACAAGACAGGACAAAGCAGCTACTTGA 2580  
Qy 570 erSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysP 590  
Db 2581 GCCCGGAGCGTGAGCTCAGAGTATGCGCAAGCTCGTACACGCGAGACATGAAAAAGT 2640  
Qy 590 heArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProT 610  
Db 2641 TCAGAAACCTGGCGTTAGAGATAAAGAAATATCAAGCAGT--GAATATGTTGGA---A 2694  
Qy 610 hrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluM 630  
Db 2695 CAAGTGATGATGGCTTAAACCTTCTGCTTCAAGTAGTGAA- - - - -GAAA 2739  
Qy 630 etGluMetGlyLysIleLysArgThrGlyGln 640  
Db 2740 TGAATAGAGGCTCAATGAACGCAATCCTCAG 2771  
RESULT 12  
ACF36552  
ID ACF36552 standard; DNA; 2820 BP.  
XX  
AC ACF36552;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Arabidopsis PERK1 receptor related protein encoding DNA.  
XX  
KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;  
KW wound; pathogen resistance; plant growth; seed production; gene; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN W02003072763-A1.  
XX  
PD 04-SEP-2003.  
XX  
PF 28-FEB-2003; 2003WO-CA000274.  
XX  
PR 28-FEB-2002; 2002CA-02373903.  
PR 28-FEB-2002; 2002US-00086464.  
XX





158	Db	ACACGCGCGTGT	CAGAACCAT	CACCAT	CCTCTCCAGATT	CACAGCTT	CCTCTCTT	ACCT	217
29	Qy	-----	-----	-----	-----	-----	-----	-----	40
218	Db	TCGATCTT	TCCTCGTAA	CAGATTC	CCACCT	CCACCT	TCGGATCT	TTCTCCACCCGTT	277
41	Qy	ProSerThr	ilePro	ThrSerPro	SerSerArg	SerThrPro	SerAla	Pro---	59
278	Db	GATTCA	ACC---	CTTCTCG	CGCCGAC	CGTCAA	ACGAATCT	CCTTCTCTCCGAA	334
60	Qy	-----	-----	-----	-----	-----	-----	-----	74
335	Db	GATTCG	GAACAC	CACCAT	CTGTCT	CCACCA	AAATGAA	TCCAATGACAA	394
75	Qy	-----	-----	-----	-----	-----	-----	-----	86
395	Db	CAAGATCT	TCAATCG	CGCTCT	CTCCAT	CTGCGCG	TGCGGA	ATGTAGACCC	454
87	Qy	GlySer	Pro	Pro---	AlaPro	ValThr	Pro	ThrArg	101
455	Db	GAATCA	CCACGTT	AACAAT	CTCTCC	AGCTCC	ACAGCAT	CAGATCT	514
102	Qy	ProSer	ValPro	GlyPro	ProSer	AsnPro	-----	SerArg	118
515	Db	CGAGTCT	CAACCAT	TAGAC	CCCTTAC	CAATCT	CCGCCA	TACAACCAT	574
119	Qy	ArgPro	ProSer	SerPro	-----	SerPro	SerPro	-----	129
575	Db	TCTCTCG	GCTAAT	CCCAAC	GCCTCC	CGCGAG	CCCAT	TCCCCACAG	634
129	Qy	-----	-----	-----	-----	-----	-----	-----	129
635	Db	CTTCTAG	TGACCT	GTGGTGT	CTCCAT	CTCTCA	CATCCCC	CTAGTAAG	694
130	Qy	-----	-----	-----	-----	-----	-----	-----	139
695	Db	CCAAAC	CAAGCA	ATGAGAT	TGGCGT	TGGCGG	CTATCA	AGGGAAG	754
140	Qy	ValGly	leAla	lleGly	GlyVal	AlaLeu	LeuVal	leValThr	159
755	Db	GTTCGT	ATGGCT	GTAGCG	GTTTCG	CAATCA	TGGCGT	TATAGCG	814
160	Qy	CysLys	LysLys	ArgArg	AspGlu	AspAla	Tyr-----	TyrVal	175
815	Db	AGAAGA	AGAAAA	AGAGA	-----	ACAT	TGATAG	TATACT	868
176	Qy	Pro	Pro	-----	-----	-----	-----	-----	177
869	Db	CATCCCA	ATTTCT	CTGTTA	ATCAGG	TTTAA	AAATCT	CACCTT	928
177	Qy	-----	-----	-----	-----	-----	-----	-----	177
929	Db	CTTCTAT	GTGCTT	GAAATCA	TCTCT	GACTAT	CTTTG	CTTTG	988
178	Qy	-----	-----	-----	-----	-----	-----	-----	194
989	Db	TACG	TCAGAT	CAAGTAA	AGGATAC	TCTCT	CTGGTCT	TAATGTT	1048
195	Qy	TrpArg	GlnGln	AsnAla	ThrPro	SerSer	AspHis	ValVal	214
1049	Db	TCACAG	CAACAA	CAATCT	CTATG	GGAACA	CAGTTAT	GTGACAG	1096
215	Qy	ProLys	AlaPro	SerPro	ProArg	GlnPro	ProPro	ProPro	234
1097	Db	-----	-----	-----	-----	-----	-----	-----	1123
235	Qy	SerSer	GlyGly	SerAsp	TyrSer	AspArg	ProVal	LeuPro	254
1124	Db	TCAAGT	GGCAC	ACCTG	AC	-----	-----	-----	1147
255	Qy	ValLeu	GlyPhe	SerLys	SerThr	PheThr	TyrGlu	LeuAla	274

Db	1148	ATACTCGGAAGTGGCCAGACTCATATTTTCAGTTTACGAAGAGCTTCTCGTAGATAACACACAAAGGC	1207
Qy	275	PhSeSerGluAlaAseLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu	294
Db	1208	TTTGCTCGCAAAACATTTCTTGGAAGGCGGATTTGGATGCTCTATATAAGGTACATTC	1267
Qy	295	ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArg	314
Db	1268	CAGGATCGTAAAGTGTGTGCGGTAAAGCAGCTTAAAGCTGGAAGTGACAAAGGTGACCGT	1327
Qy	315	GluPheGlnAlaGluValGluLeuLeuLeuSerArgValHisHisArgHisLeuValSerLeu	334
Db	1328	GAATTCAAAGCAGAGGTGTAGATCATCAGCGCGTTCATCATCGCCATTTGGTCTCTCTG	1387
Qy	335	ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn	354
Db	1388	GTGGTACTGCATTTTCAGACCAGCATAGATTCCTATCTATGAGTATGTTTCTTAATCAA	1447
Qy	355	AsnLeuGluLeuHisLysGlyGlu	363
Db	1448	ACCTTGCAGCATCATTTGCATGCTGTAGTACTTTGTTACCATTTTCGTTATAGATAAGACT	1507
Qy	364	-----GlyArgProoth	367
Db	1508	TTTTTTTTCAGCTTTCACGTGTTCAGCTCGCTTTACGCTTTTAGGAAGGGTTTGCAGT	1567
Qy	367	rMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLe	387
Db	1568	TTTAGAGTGGTCTAAGAGAGTCCGGATCGCTATAGATCAGCCAAAGGGTTGGCATATCT	1627
Qy	387	uHisGluAspCysAsn	392
Db	1628	TCAGGAAGCTGTAA - GTAATGCCTTCACATTTTCTTAGTTGTGTGCTTTGGTTATGCAC	1686
Qy	393	-----ProLys	394
Db	1687	TTTCATAGTTTAAACAGAGAAGCAAAATCATATCTCTTGTAAATCTATCTTATGACTGTAAAGTTTTCAGT	1746
Qy	394	sIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLys	414
Db	1747	AATCATTTACAGAGATATAAGTCAGCAATATTTCTTAGATGATGAATATGAGCTCA	1806
Qy	414	S-----	414
Db	1807	-GGCAATAATGAATCACTCTCTTTTCGTTAAATCTATCTTATGACTGTAAAGTTTTCAGTTA	1865
Qy	415	-----ValAlaAspPheGlyLeuAlaLysIleAl	424
Db	1866	ATGAGACTTGTCTCTGTTTTTTTGGATGTTTAGTGTGCTGATTTTGGACTCTCTAGACTCAA	1925
Qy	424	aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly	440
Db	1926	TGATACAACACAACTCATGTTTCACTCGGGTTATGGGAACCTTCGGGTGAAGCAACAT	1985
Qy	440	-----	440
Db	1986	TCATCAAAACTCTACTCCAAAAGCTGACCTTATTGATCCAAATCTCTTATGACTGAAAAAGTTTG	2045
Qy	440	-----	440
Db	2046	TTATATATATGGCTTGAGGCAACAAATTGGATCAAACTCTTTATTTGATCGTATGGCT	2105
Qy	441	-----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu	453
Db	2106	GCATGACATGTTTGTGTTAAGTACTAGCGCGGAATATGATCATCAAGTTGGAAATTTGA	2165
Qy	453	hrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg	473
Db	2166	CTGATAGATCCGATGTATTCTATTCGGGTGTGTTCTCTTAGAGCTGTGAATCTGACGCGGA	2225
Qy	473	rgProValAspAlaAsnValTyrValAspAspSerLeuValAspTrp	489
Db	2226	AACAGATTGACCAAGCTCAGCCTCTTAGGGAAGAGAGTGTGGTTGAATGGGTGAAGATCC	2285



Qy 130 -----SerSerAspGlyLeuSerThrGly-----ValVal 139  
 Db 695 CCAAAACCAAGCAATGGAGTGGCGGTGGCGGTATCATAGGGAAGACTATG 754  
 Qy 140 ValGlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeu 159  
 Db 755 GTTGGTATGGCTGTAGCCGGTTTGCATCATAGCGCTTATAGCGGTGTGTCTTAGTG 814  
 Qy 160 CysIlySlySlyAsArgAspGluGluAspAlaTyr-----TyrValPro 175  
 Db 815 AGAAGAAAGAAAAGAGA-----AACATTGATAGTATATCACTCACAGTACTTGCCA 868  
 Qy 176 ProPro----- 177  
 Db 869 CATCCCAATTTCTCTGTTAAATCAGGTTTAAAAATCTCACCTTTATCTCTCTCATCAT 928  
 Qy 177 ----- 177  
 Db 929 CTTCTATGTGCTTGAATCATCTCTGACTATCTTTGCTTTTGATGATGAGATTCTTA 988  
 Qy 178 -----ProProProGlyProIlySAlaGlyGlyProTyrGlyGlyGlnGlnGln 194  
 Db 989 TACGTCACAGATCCAGGTAAGGATACCTCTCTGCTCAATGTTCAATGTATATAACAAT 1048  
 Qy 195 TrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuProPro 214  
 Db 1049 TCACAGAACCAACAATCTCTATGCGGAACAGTTATGTGTACAGCTGGT----- 1096  
 Qy 215 ProIlySAlaProSerProArgGlnProProProProProProProProPheMetSer 234  
 Db 1097 -----GGTGTTATCTCATCATCA-----ATGCCAA 1123  
 Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProProSerProGlyLeu 254  
 Db 1124 TCAAGTGCACACCTGAC-----TCGTCT 1147  
 Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274  
 Db 1148 ATACTCGGAAGTGGCCAGACTCAVTTTCAGTTTACGAAGAGCTTGTGTGATTAACAAGG 1207  
 Qy 275 PheSerGluAlaAsnLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu 294  
 Db 1208 TTGTCTCCAAAACATTTCTTGGAGAGCGGATTTGGATGTGTCTATAAAGGTACATTG 1267  
 Qy 295 ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyArg 314  
 Db 1268 CAGGATGGTAAAGTTGTTGCGTTAAAGCAGCTTAAAGCTGGAAGTGGACAAGGTGACCGT 1327  
 Qy 315 GluPheGlnAlaGluValGluIleSerArgValHisHisArgHisLeuValSerLeu 334  
 Db 1328 GAATTCAAAGCAGAGGTTGAGATCATCAGCCGCTTCATCATCGCCATTTGCTCTCTG 1387  
 Qy 335 ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn 354  
 Db 1388 GTTGGTACTGCATTTTCAGACAGCATAGATTGCTTATCTATGATGATGTTCTTAATCAA 1447  
 Qy 355 AsnLeuGluLeuHisLeuHisGlyGlu----- 363  
 Db 1448 ACCTTGGAGCATCATTTGCTGAGTGAGTGACTTGTACCATTTCTCGTTATAGATAAGACT 1507  
 Qy 364 -----GlyArgProTh 367  
 Db 1508 TTTTTTAGCTTTAGTGTAGACTGCTCGCTTTACGCTTTAGGAAAGGTTTGGCCAGT 1567  
 Qy 367 rMetGlnTrpSerThrArgLeuValIleAlaLeuGlySerAlaLysGlyLeuSerTyrLe 387  
 Db 1568 TTTAGATGGTCTTAAGAGAGTCCGATCGCTATAGGATCAGCCAAAGGTTGGCATATCT 1627  
 Qy 387 uHisGluAspCysAsn----- 392  
 Db 1628 TCACGAAGACTGTAA-GTAATGCCCTTCACATTTCTTAGTTGTGTGCTTTGGTTATGCAC 1686

Qy 393 -----ProLy 394  
 Db 1687 TTCTAGTGTAAACAGAAAGCAAAATCATATCTTGTGTTTATTTTTCAGGTCTACCGAA 1746  
 Qy 394 sIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLy 414  
 Db 1747 AATCATTTCCAGAGATATAAGTCAGCAATATTTCTTAGATGATGATGATGAGCTCA 1806  
 Qy 414 s----- 414  
 Db 1807 -GGCAATAATGAATCCTCTCTTTTCGTTAAATCTATCTTATGACTGTAAAGTTTAGTTA 1865  
 Qy 415 -----ValAlaAspPheGlyLeuAlaLysIleAl 424  
 Db 1866 ATGAGACTGTTGTTGTTTTTGGATGTTAGTTGCTGATTTTGGACTTGTAGACTCAA 1925  
 Qy 424 aserAspThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly 440  
 Db 1926 TGATACACACAACTCATGTTTCAACTCGGGTTATGGGAACCTTCGGGTAAAGCAACAT 1985  
 Qy 440 ----- 440  
 Db 1986 TCATCACAAACTCTACTCTCCAAACTCGACCTTATTGATCCAATGCTCATGAAAGTTTG 2045  
 Qy 440 ----- 440  
 Db 2046 TTATATATGGCTTGAGCAACAAATGGATCAAACTGAATCTTTATTGATCGGTATGGCT 2105  
 Qy 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 453  
 Db 2106 GCATGACATGTTTGTGTTAAGGTACCTAGCGCGGAATATGCATCAAGTGGNAATTTGA 2165  
 Qy 453 hrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 473  
 Db 2166 CTGATAGATCCGATGTTCTCTATTCGGGGTTGTTCTCTTAGAGTGTCTACTGACGGA 2225  
 Qy 473 xgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp----- 489  
 Db 2226 AACCACTTACAGACTCAGCTCCTAGGAGAGAGAGTTGGTTGAAATGGGTAAAGATCC 2285  
 Qy 489 ----- 489  
 Db 2286 AACTTTCAACATTTCTTAATAATAGTAGTGGCCCTAGTATATATATAGTACTTA 2345  
 Qy 490 -----AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlu 504  
 Db 2346 TAAATGAACCTCACAGGCGCGCCGCTCTCTCAAGCCATTGAGACCGGAGATTTAAGC 2405  
 Qy 505 GlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetVal 524  
 Db 2406 GAATCTGATTGATACACCGCTTGAAAAGCGTTATGTGGAGCATGAAGTCTTCAGATGATC 2465  
 Qy 525 AlaCysAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGln--- 543  
 Db 2466 GAGACAGCGCTGCATGTGTTAGACATTTCTGGTCCAAAACGCCCAACGATGTTTCAGGTA 2525  
 Qy 543 ----- 543  
 Db 2526 ATTCGTACTAACCAAAAGTCCAAAGCTCCCATATATATAGTAACAAAGTATTTCTCACATCT 2585  
 Qy 544 -----IleValArgAlaLeuGluGlyAsnValSerLeu 554  
 Db 2586 GAAACCTTATCTACTCTTCGAAATAAGGTTGTGAGACATTTGGACTCGGACGGAGACTCG 2645  
 Qy 555 SerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerTyrGlyGly 574  
 Db 2646 GGATATATTACCAACGGATCAAAATTTGGSCAA----- 2678  
 Qy 575 SerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysPheArgLysMetAla 594  
 Db 2679 AGCAAACTTATGACTCAGGGCAATACAATGAAGACATTTATGAAATTCAGGAAATGGCG 2738  
 Qy 595 LeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGly 614





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 24, 2004, 20:12:36 ; Search time 3601 Seconds  
(without alignments)  
5365.407 Million cell updates/sec

Title: US-10-086-464-2  
Perfect score: 3453  
Sequence: 1 MSSAPSGTGTGPPSPNST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 segs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=x1p  
-Q=/cg2\_1/USPTO\_spool\_p/US10086464/runat\_23042004\_083015\_5625/app\_query.fasta\_1.839  
-DB=EST -QFMT=fastcap -SUFFIX=rbt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10086464\_QCGN\_1\_1\_4237 @runat\_23042004\_083015\_5625 -NCPU=6 -ICPU=3  
-NO\_MMAB -LARGESQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1092.5	31.6	1084	12	BP184771	BP184771 BP184771
2	1090.5	31.6	824	14	CD435142	CD435142 EL01N0355
3	1086	31.5	788	14	CF436437	CF436437 EST672782
4	1071	31.0	789	14	CF436655	CF436655 EST673000
5	1069	31.0	759	12	BM408099	BM408099 EST582426
6	1066	30.9	785	14	CF436122	CF436122 EST672467
7	1065	30.8	757	12	BG596561	BG596561 EST495239
8	1063	30.8	731	14	CB655196	CB655196 OSJNEC08F
9	1047	30.3	692	12	BG441204	BG441204 GA_Ea001
10	1030	29.8	723	14	CD839231	CD839231 RF02_1141
11	1027.5	29.8	1053	11	AY108241	AY108241 Zea mays
12	1025	29.7	1016	11	AY108243	AY108243 Zea mays
13	1018	29.5	666	12	BM358715	BM358715 GA_Ea001
14	1013	29.3	695	13	CA072174	CA072174 SCCAM100
15	983	28.5	715	13	BU099573	BU099573 WHE3309_A
16	961	27.8	770	13	BO999193	BO999193 QG21G04
17	931	27.0	611	14	CA237156	CA237156 SCWCF1500
c 18	931	27.0	859	29	CC725849	CC725849 OGAAB66TC
c 19	929.5	26.9	809	14	CA765135	CA765135 AF53-Rpf
c 20	928	26.7	850	29	CC668754	CC668754 OGAAB66TM
c 21	923.5	26.7	769	14	CK283399	CK283399 EST746121
c 22	921	26.7	618	14	CA269355	CA269355 SCBGR7307
c 23	921	26.7	651	14	CF478389	CF478389 RTW3_18
c 24	920	26.6	673	14	CA164704	CA164704 SCBGR2312
c 25	919.5	26.6	938	29	CG333846	CG333846 OGAAB66TC
c 26	917.5	26.6	949	29	CG436431	CG436431 OGVGH77TV
c 27	912	26.4	679	13	BQ404121	BQ404121 GA_EG006
c 28	904.5	26.2	674	14	CA298046	CA298046 SCWCF1500
c 29	903	26.2	595	14	CF015663	CF015663 QBL8b11.X
c 30	903	26.2	637	13	BQ34241	BQ34241 1091014D0
c 31	903	26.2	645	13	BQ506869	BQ506869 EST614284
c 32	897.5	26.0	899	14	CA096498	CA096498 UB16CPE10
c 33	895.5	25.9	682	13	CA095337	CA095337 SCCCL1500
c 34	895	25.9	641	14	CA248316	CA248316 SCWCF1500
c 35	895	25.9	725	29	CG450877	CG450877 OGAAB66TC
c 36	893.5	25.9	803	14	CF243440	CF243440 3530_121
c 37	891	25.8	646	13	BQ240617	BQ240617 TAE05015B
c 38	891	25.8	652	13	BQ849683	BQ849683 QGB10120
c 39	891	25.8	666	13	BU499754	BU499754 946178G10
c 40	891	25.8	691	14	CA237152	CA237152 SCWCF1500
c 41	889	25.7	626	13	BQ582873	BQ582873 E012278-0
c 42	888.5	25.7	593	10	BF176907	BF176907 EMI_4_B10
c 43	887	25.7	622	14	CA999610	CA999610 S345P_H09
c 44	887	25.7	673	13	BQ410602	BQ410602 GA_EG003
c 45	886.5	25.7	573	9	AV551753	AV551753 AV551753

ALIGNMENTS

RESULT 1  
BP184771  
LOCUS  
DEFINITION BP184771 pns rice panicle cDNA, germ cell generating stage Oryza sativa (japonica cultivar-group) cDNA, mRNA sequence.  
ACCESSION BP184771  
VERSION BP184771.1  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.  
1084 bp mRNA linear EST 17-JUL-2003

```

REFERENCE
AUTHORS      1. (bases 1 to 1084)
TITLE        Moriguchi,K., Ito,Y., Yamazaki,Y. and Kurata,N.
JOURNAL      Finding of various plant nuclear proteins using yeast nuclear
COMMENT      transportation trap system - a proteomal approach
              Unpublished (2003)
              Contact: Kazuki Moriguchi
              Plant Genetics
              National Institute of Genetics
              Yata 1111, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-55-981-6872
              Fax: 81-55-981-6879
              Email: kmoriguc@lab.nig.ac.jp
              cDNA clone obtained from nuclear transportation trap system
              encoding a protein similar to Oryza sativa (japonica
              cultivar-group) putative receptor protein kinase PERK1.
              Location/Qualifiers
                1..1084
                  /organism="Oryza sativa (japonica cultivar-group)"
                  /mol_type="mRNA"
                  /db_xref="taxon:39947"
                  /tissue_type="panicle"
                  /dev_stage="germ cell generating stage"
                  /clone_lib="pNS rice panicle cDNA, germ cell generating
                  stage"

FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.:      4.98e-26      Length:      1084
Score:          1092.50      Matches:     212
Percent Similarity: 82.88%      Conservative: 30
Best Local Similarity: 72.60%      Mismatches: 43
Query Match:    31.64%      Indels:      7
                  12      Gaps:      3

US-10-086-464-2 (1-647) x BP184771 (1-1084)

Qy      360  LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db      2   TTGCACGGAAGAGCGCCACCAATGAGTGCGCCCAAGACTAAAGATTGCTTTGGGA 61

Qy      380  SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db      62  GCTGCAGAAAGGTTTAGCTTAATCTTCATGAGAGCTGCCATCTTAAGATCATTCATCGTAT 121

Qy      400  IleLysAlaSerAsnIleLeuLysPheLysPheGluAlaLysValAlaAspPheGly 419
Db      122  ATTAAGGGCTCAAAACATCTTCTTGATTGAATTTGAATCTAAGGTTGCTGATTGGA 181

Qy      420  LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db      182  CTTGCTAAGTTCCACGAGTGATTAATACACTCATGTTTCGCAAGGTAATGGGCACTTTT 241

Qy      440  GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db      242  GGATATCTAGCACACAGAGTACGATCTTCTGCAAGCTCACTGGAATAATCAGATGCTCTTC 301

Qy      460  SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db      302  TCCTATGGAGTTATGCTTCTTGAGTTAATAACTGCTGCGCCAGTTGTATACAGTCAA 361

Qy      480  ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db      362  ACATATATGGATAGACAGCTTGTTGTTACCTGGGCAAGGCTTTTACTGATCGAAGCACTTGG 421

Qy      500  GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db      422  AATGGTAACACGAGGAGTTAGTAGATCTCTGGCTTGGGAAGGATTTCAATCCCAATGAG 481

Qy      520  MetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgPro 539
Db      482  ATGGCAGAAATGATTGCTTGTGACATGCTGATCGCTACGCCAATTCGCTCGTCGCGCCCA 541

Qy      540  ArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGlu 559

```

```

Db      542  CGCATGACGACAGTGTTCGGGCTTTGGAAGGTGACGTGCTTTGGAGGATCTTAATGAA 601
Qy      560  GlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAsp 579
Db      602  GGTGTTCCGGCTGGTTCACAGCCGCTATTATTGGATCGTAC---AGCAGCTCTGACTATGAT 658
Qy      580  SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGlu 599
Db      659  TCTGGCCAATACACAGGAGCATGAAGAAGTTTCAGGAAGATGGCTTTT---ACCAACAAT 715
Qy      600  TyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGly 619
Db      716  AATGATACGAGCAGCTCAATACAGCGCACCAACACGAGGATATGCCAGATACCTCTGCA 775
Qy      620  SerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly 639
Db      776  TCAAGCAGCGAGGCGCCAAACAAACCCAGAAAGTCAGAGCAGGACAAACCAAGAGAGCGGC 835
Qy      640  -----GlnGlyTyrSerGlyProSer 646
Db      836  TACAGTGGCTACAGCTCAGGATACGCGGAGCCTCA 871

RESULT 2
CD435142
LOCUS      CD435142      824 bp      mRNA      linear      EST 03-JUN-2003
DEFINITION E01N035D03.b Endosperm_3 Zea mays cDNA, mRNA sequence.
ACCESSION  CD435142
VERSION    CD435142.1  GI:31350785
KEYWORDS   EST.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (bases 1 to 824)
AUTHORS   Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
            Messing,J.
TITLE      Sequencing of the maize endosperm ESTs
JOURNAL    Unpublished (2002)
COMMENT    Dr. Joachim Messing's lab
            Wakeman Institute, Rutgers University
            190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
            Tel: 732-445-3801
            Fax: 732-445-5735
            Email: jlai@wakeman.rutgers.edu
            Seq primer: T3.
FEATURES   Location/Qualifiers
            source          1..824
                           /organism="Zea mays"
                           /mol_type="mRNA"
                           /db_xref="taxon:4577"
                           /tissue_type="Endosperm of 7-23DAP"
                           /clone_lib="Endosperm_3"
                           /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                           XhoI"
ORIGIN
Alignment Scores:
Pred. No.:      4.54e-26      Length:      824
Score:          1090.50      Matches:     206
Percent Similarity: 84.12%      Conservative: 27
Best Local Similarity: 74.37%      Mismatches: 41
Query Match:    31.58%      Indels:      3
                  14      Gaps:      2

US-10-086-464-2 (1-647) x CD435142 (1-824)

Qy      334  LeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsn 353
Db      2   TTGGTTGGCTATTGCAITTTCTTGGAGGACCTTGCTGCTTGTCTATGATGATTTGTCCCAAT 61

```

Qy 354 AenAenLeuGluLeuHisLeuHisGlyGluArgProThrMetGluTrpSerThrArg 373  
 Db 62 AACACATGGAAATCCCACTTACATCGCGAAGGTGACCACTATGGAGTGGCCCTGCTAGA 121  
 Qy 374 LeuYsileAlaLeuGlySerAlaLysGlyLeuSerTyLeuHisGluAAspCysAAsnPro 393  
 Db 122 TTAAGATCAGTTGGGTGCTGCCAAGGGTTAGCTTATCTTATCTTCAAGAGCTGCATCCA 181  
 Qy 394 LysileileHisArgApilleyAlaSerAenileLeuileAAspPheLysPheGluAla 413  
 Db 182 AAGATCATCCATCGTGACATAAGGCATCAACATTTCTTGACTTCCCAATTTGAAGCT 241  
 Qy 414 LysValAlaAAspPheGlyLeuAlaLysileAlaSerAAspThrAenThrHisValSerThr 433  
 Db 242 AAGGTGCTGATTTGGCTTGCAGTTCTCAAGTTCACTACTGATACACACCCATGTTTCGACA 301  
 Qy 434 ArgValMetGlyThrPheGlyTyLeuAlaProGluTyAlaAlaSerGlyLysLeuThr 453  
 Db 302 AGAGTAATGGGCACCTTTGGGTATTTGGCACCTGAGTATGCATCTTCTGGCAAGCTAAACA 361  
 Qy 454 GluLysSerAAspValPheSerPheGlyValValLeuLeuGluLeuileThrGlyArgArg 473  
 Db 362 GAAAAATCCGATGATTTTCTCGAGTCATGCTTCTTGAGCTTATTTACTGGGCGCGA 421  
 Qy 474 ProValAAspAlaAAsnAAsnValTyValAAspSerLeuValAAspTrpAlaAAspProLeu 493  
 Db 422 CCAGTTGACACAAACCAACATATGATGACAGCTTGTGTGACTGGCAAGGCCATTA 481  
 Qy 494 LeuAenArgAlaSerGluGlnGlyAAspPheGluGlyLeuAlaAAspAlaLysMetAAsn 513  
 Db 482 CTGATGCGAGCACCTTGAGGATGGTGAATATGATCTTTAGTGGATCTCTCGGCTGGGAAG 541  
 Qy 514 GlyTyAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAAspValAAspHis 533  
 Db 542 GACTTCAATCTAATAGATGGCAAGATGATAGCTGTGCGAGCTGATGTGTACGGCAT 601  
 Qy 534 SerAlaArgArgArgProArgMetSerGlnileValArgAlaLeuGluGlyAAsnValSer 553  
 Db 602 TCTGACAGTCTCGGCCACGATGATGATGATGATGATGATGATGATGATGATGATGAT 661  
 Qy 554 LeuSerAAspLeuAenGluGlyMetArgProGlyGlnSerAenValTySerSerTyGly 573  
 Db 662 TTGGAGGACCTTAATGAAGGTGTTTGGCTGGCCATAGCCGCTTCTTTGGGTCTATAC- 718  
 Qy 574 GlySerThrAAspTyAspSerSerGlnTyAAsnGluAAspMetLysLysPheArgLysMet 593  
 Db 719 AGCAGCTCCGATACGATTTCTGGCAGCTACCAACGAGGACATGAAGAGATTTCAAGAGATG 778  
 Qy 594 AlaLeuGlyThrGlnGluTyAsnAlaThrGlyGluTySerAAsnProThr 610  
 Db 779 GCATTCACACACAC- - - - -TATACGAGCAGCCCAATACAGCGGCCACCC 823  
 RESULT 3  
 LOCUS CF436437  
 DEFINITION EST672782 normalized cDNA library of onion Allium cepa cDNA clone  
 ACACM44, mRNA sequence.  
 ACCESSION CF436437  
 VERSION CF436437.1 GI:34459127  
 KEYWORDS EST.  
 SOURCE Allium cepa (onion)  
 ORGANISM Allium cepa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;  
 Allium.  
 1 (bases 1 to 788)  
 HAVEMJ, Cheung, F., Van Aken, S., Utterback, T. and Town, C.D.  
 TITLE Expressed Sequence Tags from a normalized library of mixed onion  
 tissues (Allium cepa)  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Haver MJ  
 Department of Horticulture

USDA-ARS and University of Wisconsin  
 1575 Linden Drive, Madison, WI 53706, USA  
 Tel: 608-262-1830  
 Fax: 608-262-4743  
 Email: mjhavey@facstaff.wisc.edu  
 TIGR sequence name ACACM44TR. For more information:  
 http://haveylab.hort.wisc.edu  
 Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES  
 source  
 1..788  
 /organism="Allium cepa"  
 /mol\_type="mRNA"  
 /cultivar="Red Creole (bulbs), unknown(callus), Ebano &  
 Texas Legend(roots)"  
 /db\_xref="taxon:4679"  
 /clone="ACACM44"  
 /tissue\_type="Callus, roots, and young bulbs"  
 /clone\_lib="normalized cDNA library of onion"  
 /notes="vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:  
 EcorV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA  
 from callus, roots, and young bulbs were combined to  
 synthesize the library. Normalization to enrich for  
 low-copy transcripts was performed by proprietary  
 techniques of Invitrogen."

Alignment Scores:  
 Pred. No.: 6,07e-26 Length: 788  
 Score: 1086.00 Matches: 204  
 Percent Similarity: 87.79% Conservative: 26  
 Best Local Similarity: 77.86% Mismatches: 32  
 Query Match: 31.45% Indels: 0  
 DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x CF436437 (1-788)

## ORIGIN

Qy 260 LysSerThrPheThrTyLeuGluLeuAlaArgAlaThrAenGlyPheSerGluAlaAAsn 279  
 Db 2 AAGAGCACTTTCACATATGAAGAATGGCGATAGCTACAAATGGTTTCCGACTATAAT 61  
 Qy 280 LeuLeuGlyGlnGlyPheGlyTyLeuValHisLysGlyValLeuProSerGlyLysGlu 299  
 Db 62 CTTCTTGGGCAAGGTGGATTTGGATATGTGCACAAAGAGTACTTCCAAACGGTAAAGAA 121  
 Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319  
 Db 122 GTGGCTATCAAGCAGTTGAAGCCGGAAGCGGAGCGGTGAGTTTCAAGCAGAG 181  
 Qy 320 ValGluileileSerArgValHisArgHisLeuValSerLeuValGlyTyCysile 339  
 Db 182 GTTGAATATATCATGTCGAGTGCATCATAGGCATTTGGTTTCTTTAGTTGGCTATTGCATT 241  
 Qy 340 AlaGlyAlaLysArgLeuLeuValTyLeuPheValProAenAenAenLeuLeuHis 359  
 Db 242 TCAGGAGATCATAGATTGCTGTGTATGATATGTTCTTAATAAACCCTTGATTTCCAT 301  
 Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysileAlaLeuGly 379  
 Db 302 TTGCATGGAANAAGATGTTCCACTGATGGATTGGCAACGGGTAAANAATGCTTTGGGT 361  
 Qy 380 SerAlaLysGlyLeuSerTyLeuHisGluAAspCysAAsnProLysileileHisArgAAsp 399  
 Db 362 TCTGCCAAGGGTTGGCATATCTCCATGAAGATTGTTCATCCCAAAATTTATTCATCGTAT 421  
 Qy 400 IleLysAlaSerAenileLeuileAAspPheLysPheGluAlaLysValAlaAAspPheGly 419  
 Db 422 ATTAAGCAGCAAAATATCTTCTTGATGATGAACCTTTGAGGCTAAGGTTGCAGATTTTGGC 481  
 Qy 420 LeuAlaLysileAlaSerAAspThrAenThrHisValSerThrArgValMetGlyThrPhe 439  
 Db 482 CTTGCAAAAGTTTCACTGTTGAAATAAACACCCATGTTTCTACAGAGTCATGGGAACATTT 541  
 Qy 440 GlyTyLeuAlaProGluTyAlaAlaSerGlyLysLeuThrGluLysSerAAspValPhe 459

```

|||||GGGTATCGGCACCAATATGATCCCTCTGTGTAACTAAGTGAATAATCAGATGCTCTTT 601
Db
460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Qy
602 TCAITTTGAGTCATGCTTTTGGAGCTGATTACTGGAAAGCGCAGTTGATGCAACCAA 661
Db
480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Qy
662 ACTTATACAGATGATAGTTTGGTCGATTGGGCAAGACCATTGCTGCAGAGGCTATGGAG 721
Db
500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Qy
722 CATGATGATTATGATGAGCTCATTCACAGAAGGCTTAGGAGATAAATAACAACATGACGAA 781
Db
520 MetAla 521
Qy
782 ATGGCA 787
Db

RESULT 4
CF436655
LOCUS EST673000 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACACK26 mRNA sequence.
ACCESSION CF436655
VERSION CF436655.1 GI:34459345
KEYWORDS EST.
SOURCE Allium cepa (onion)
ORGANISM Allium cepa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
Allium.
REFERENCE 1 (bases 1 to 789)
AUTHORS Havey,M.J., Cheung,F., Van Aken,S., Utterback,T. and Town,C.D.
TITLE Expressed Sequence Tags from a normalized library of mixed onion
tissues (Allium cepa)
JOURNAL Unpublished (2003)
COMMENT Contact: Havey MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhavey@facstaff.wisc.edu
TIGR sequence name ACACK26TR. For more information:
http://haveylab.hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.
FEATURES
source
location/Qualifiers
1..789
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole (bulbs), unknown (callus), Ebano &
Texas Legend (roots)"
/db_xref="taxon:4679"
/clone="ACACK26"
/tissue_type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/notes="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:
EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
from callus, roots, and young bulbs were combined to
synthesize the library. Normalization to enrich for
low-copy transcripts was performed by proprietary
techniques of Invitrogen."

ORIGIN
Alignment Scores:
Pred. No.: 1,82e-25 Length: 789
Score: 1071.00 Matches: 203
Percent Similarity: 86.64% Conservative: 24
Best Local Similarity: 77.48% Mismatches: 35
Query Match: 31.02% Indels: 0
DB: 14 Gaps: 0

```

```

US-10-086-464-2 (1-647) x CF436655 (1-789)
Qy 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
Db 2 AAGAGCACTTTCACATATGAAGAATTGGCGATAGTACAAATGGGTTTCCGACTATAAT 61
Qy 280 LeuLeuGlyGlnGlyPheGlyTyrValHisGlyValLeuProSerGlyValGlu 299
Db 62 CTTCTTGGCAAGGTGGATTGGATATGTGCACAAAGGAGTACTTCCAAACGTAAGAA 121
Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgPheGlnAlaGlu 319
Db 122 GTGGCTATCAACGAGTTGAAGCCGGAAGCGCAAGGGGAGCGTGAGTTTCAAGCAGAG 181
Qy 320 ValGluIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIle 339
Db 182 GTTGAATAATTATCAGTCGAGTGCATCATGAGCATTTGGTCTTCTTTAGTTGGCTATTGCATT 241
Qy 340 AlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHis 359
Db 242 TCAGGAGATCATGATGCTTGTCTATGAATATGTTCTTAATAAAACCCCTTGATTCAT 301
Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 302 TTGCATGGAAAGATGTTCCACCTATGATGGCAACGGGTTAAAAATTTGCTTTGGGT 361
Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 362 TCTGCCAAGGGTTGGCATATCTCCATGAAGATTGTTCATCCCAAAATTAATTTCATCGTAT 421
Qy 400 IleLysAlaSerAsnIleLeuIleAspPheIleGluAlaLysValAlaAspPheGly 419
Db 422 ATTAAGACAGCAAAATATTTCTTTGAGATGAACATTTGAGGCTAAGGTTGCAGATTTTGGC 481
Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 482 CTTGCAAGTTCACTGTTGAAATAACACCCATGTTTCTACACGAGTCATGGGAACATTT 541
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 542 GGGTATCTGGCACCAGAAATATGCATCTCTGGTAAACTAAGTGTATAAATCAGATGTCTTT 601
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 602 TCATTTGGAGTCATGCTTTTGGAGCTGATTACTGGAAGCGCAGCATTTGATGCAACCCAA 661
Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 662 ACTTATACAGATGATAGTTTGGTCGATTGGCAAGACCATTGCTGCAGAGGCTATGGAG 721
Qy 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db 722 CATGATGATTATGATGAGCTCATTCACAGAAGGCTAGGAGATAAATACCCCATGACGA 781
Qy 520 MetAla 521
Db 782 ATGGCA 787

RESULT 5
BM408099 759 bp mRNA linear EST 10-MAR-2003
LOCUS EST582426 potato roots Solanum tuberosum cDNA clone CP033021 5'
DEFINITION end, mRNA sequence.
ACCESSION BM408099
VERSION BM408099.1 GI:18259729
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 759)
AUTHORS van der Hoeven,R., Sun,H., Karamycheva,S.A., Tsai,J., Van Aken,S.,

```

Utterback, T., Chiemingo, A., Bougri, O., Buell, C.R., Romning, C., Tanksley, S. and Baker, B.  
Generation of ESTs from potato roots  
Unpublished (2001)  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: T3.

# FEATURES

Location/Qualifiers  
1..759  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="CPRO33021"  
/tissue\_type="roots"  
/dev\_stage="in vitro grown stem cuttings"  
/lab\_host="SOLR"  
/clone\_lib="potato roots"  
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

## ORIGIN

Alignment Scores:  
Pred. No.: 2,04e-25 Length: 759  
Score: 1069.00 Matches: 200  
Percent Similarity: 89.29% Conservative: 25  
Best Local Similarity: 79.37% Mismatches: 27  
Query Match: 30.96% Indels: 0  
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BM408099 (1-759)

Qy 312 GlyGluArgGluPheGlnAlaGluValGluLeuSerArgValHisArgHisLeu 331  
Db 2 GGGGAACGTGAATTCAGCGGAAGTTGAGATTATTAGCCGAGTACATCAAGCATCTT 61  
Qy 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuValTyrGluPheVal 351  
Db 62 GTGTCTCTTGTGGATACTGCATTTACTGGGGCTCAGAGACTGCTTGTATTGAGTTTGT 121  
Qy 352 ProAenAenLeuGluLeuHisGlyGluGlyArgProThrMetGluTrpSer 371  
Db 122 CCAACAACTACTTTGGAAATTCATTACAGGAAGGAGGCTCTCTTGGATTGGCCA 181  
Qy 372 ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys 391  
Db 182 ATACGGCTAAAGATTGCTCTAGGCTCAGCTAAAGGAGCTGGCATATCTGCATGAAGACTGC 241  
Qy 392 AsnProLysIleIleHisArgAspIleLysAlaSerAenIleLeuIleAspPheLysPhe 411  
Db 242 CAACCGAAATCAATTCACCGTGATATCAAGGAGCTAATATATCTATCGACTTTAAATTT 301  
Qy 412 GluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAenThrHisVal 431  
Db 302 GAGGCTAAGGTGCTGATTTTGGACTTGGCAAGCTAACTTCTGATGTTAATCTCATGTC 361  
Qy 432 SerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLys 451  
Db 362 TCCACGAGGTGATGGGAACCTTTGGGTATTGGCTCCAGATATGCTTCTCTGGAAG 421  
Qy 452 LeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGly 471  
Db 422 CTTACAGACAAAGTCAGACGATTTCTCTTCTTGGTGAATGCTTCTTGATTTGATAACTGGA 481

Qy 472 ArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArg 491  
Db 482 CGTCGGCTGTGTGACTCTACTCAATCATACATCGAAGATAGTTTGGTGGACTGGCAGCT 541  
Qy 492 ProLeuLeuAenArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMet 511  
Db 542 CCATTACTCACAGAGCTTTAGAGAGATGAAGAAGTTTGATACCCCTTGTGTGTCGGCTA 601  
Qy 512 AsnAenGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysVal 531  
Db 602 GAAATGATTATTAACCATATAGATGCTCGCATGTTGCTTGTGCTGCTGTGTG 661  
Qy 532 ArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsn 551  
Db 662 CGTCATTTCAGCAAGCGTAGACCAAGATGACACAGGTTCTCCGAGCTTGGGAAGGAGAT 721  
Qy 552 ValSerLeuSerAspLeuAenGluGlyMetArgPro 563  
Db 722 GTCTCATTTACAGACCTTTACGAAGGAGATTAAACCT 757

## RESULT 6

LOCUS CF436122 785 bp mRNA linear EST 04-SEP-2003  
DEFINITION EST672467 normalized cDNA library of onion Allium cepa cDNA clone ACACJ26, mRNA sequence.

ACCESSION CF436122

VERSION CF436122.1 GI:34458812

KEYWORDS EST.

SOURCE Allium cepa (onion)

ORGANISM Allium cepa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae; Allium.

REFERENCE 1 (Bases 1 to 785)

AUTHORS Havey, M.J., Cheung, F., Van Aken, S., Utterback, T. and Town, C.D.

TITLE Expressed Sequence Tags from a normalized library of mixed onion tissues (Allium cepa)

JOURNAL Unpublished (2003)

COMMENT Contact: Havey MJ

Department of Horticulture

USDA-ARS and University of Wisconsin

1575 Linden Drive, Madison, WI 53706, USA

Tel: 608-262-1830

Fax: 608-262-4743

Email: mjhavey@facstaff.wisc.edu

TIGR sequence name ACACJ26TR. For more information:

http://haveylab.hort.wisc.edu

Seq primer: CAG GAA ACA GCT ATG ACC.

FEATURES Location/Qualifiers

source

1..785

/organism="Allium cepa"

/mol\_type="mRNA"

/cultivar="Red Creole (bulbs), unknown (callus), Ebano &

Texas Legend (roots)"

/db\_xref="taxon:4679"

/clone="ACACJ26"

/tissue\_type="Callus, roots, and young bulbs"

/clone\_lib="normalized cDNA library of onion"

/notes="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:

EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA

from callus, roots, and young bulbs were combined to

synthesize the library. Normalization to enrich for

low-copy transcripts was performed by proprietary

techniques of Invitrogen."

## ORIGIN

Alignment Scores:

Pred. No.: 2,61e-25 Length: 785

Score: 1066.00 Matches: 204

Percent Similarity: 86.64% Conservative: 23

Best Local Similarity: 77.86% Mismatches: 34

Query Match: 30.87% Indels: 1

DB: 14 Gaps: 0

```
US-10-086-464-2 (1-647) x CF436122 (1-785)

QY 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
DB 2 AAGAGCACTTTCACATATGAAGATTGGCGATAGCTACAAATGGTTTCCGACTATAAT 61
QY 280 LeuLeuGlyGlnGlyGlyPheGlyTyrValHisGlyValLeuProSerGlyLysGlu 299
DB 62 CTTCCTGGCAAGGTGGATTGGATATGTGCACAAAGGAGTACTTCCAAACGGTAAAGAA 121
QY 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
DB 122 GTGGCTATCAAGCAGTTGAAGCCGGAAGCGAAGCGGAGCGTGAGTTTCAAGCAGAG 181
QY 320 ValGluIleIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIle 339
DB 182 GTTGAATATACAGTCGAGTCATCATAGGCATTTGGTTCTTTAGTTGGCTATTGCATT 241
QY 340 AlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHis 359
DB 242 TCAGGAGATCATAGATCTTGTCTATGAATATGTTCCTAATAAAACCCCTTGAGTTCAT 301
QY 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
DB 302 TTGCATGGAAAGATGTTCCACTATGATTTGGCCACCGCGTTAAATAATTGTCTTTGGGT 361
QY 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
DB 362 TCTGCCAAGGGTGGCATATCTCCATCAAGATTGTCTATCCCAAAATTTATTCATCGTAT 421
QY 400 IleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGly 419
DB 422 ATTAAGCAGCAAAATATTCTCTTGAGATGAACTTTGAGGCTAAGGTTGCGAGATTTGGC 481
QY 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
DB 482 CTGTGCAAGTTCACGTGGTGAATAAACACCCATGTTTCTACACGAGTCATGGGAACATT 541
QY 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
DB 542 GGGTATCTGGCCAGCATATGATCTCTCTGTGTAACCTAAGTGATAAATCATGATGCTTT 601
QY 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
DB 602 TCATTGGAGTCATGCTTTTGGAGCTGATTACTTGGAAAGCGACCATGTTGATGCAACCCAA 661
QY 480 ValTyrValAspAspSer-LeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerG1 499
DB 662 ACTTATACAGATGATAGATTTTGGTCGATGGCGAAGACCATTTGTCGAGAAGGCTATGGA 721
QY 499 uGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluG1 519
DB 722 GCATGATGATTATGATGAGCTTCATTGACAGAAAGCGCTAGGAGATAATACAACCATGACA 781
QY 519 uMet 520
DB 782 AATG 785

RESULT 7
LOCUS BG596561
DEFINITION EST495239 cSTS Solanum tuberosum cDNA clone cSTS15A23 5' sequence,
mRNA sequence.
ACCESSION BG596561
VERSION BG596561.1 GI:13614701
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 757)

AUTHORS van der Hoeven,R., Berzerides,J., Sun,H., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R.
FEATURES
source
1..757
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS15A23"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/clone_lib="cSTS"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
ORIGIN
Alignment Scores: 2.72e-25 Length: 757
Pred. No.: 1065.00 Matches: 200
Score: 89.16% Conservative: 22
Percent Similarity: 80.32% Mismatches: 27
Best Local Similarity: 30.84% Indels: 0
Query Match: 12 Gaps: 0
DB:
US-10-086-464-2 (1-647) x BG596561 (1-757)
QY 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
DB 1 GCAGTTAAACAGCTTAAGCTGGAGGTGGAGTGGACAGAGGGAGACGTGAATTTTCAGCCGAAGTT 60
QY 321 GluIleIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
DB 61 GAGATTATTAGCCGAGTACATCACAGCATCTTGTGTCTCTTTGGATACTGTCATTACT 120
QY 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360
DB 121 GGGGCTCAGAGACTGCTTGTATTATGAGTTTGTCCAAACAATACTTTTGGAAATTTTCATTTA 180
QY 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
DB 181 CACGAAAGGGAAGCGCTCTCTTGGATTGGCAATACGGCTAAGATTTGCTCTAGGGTCA 240
QY 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
DB 241 GCTAAGGACTGGCATATCTGCATGAAGACTGCCAACCGAAATCAATTCACCGTGATATC 300
QY 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
DB 301 AAGGCAGCTAATATACATTCGACTTTTAATTTTTCAGGCTAAGGTTGCTGATTTTGGACTT 360
QY 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
DB 361 GCCAAGCTAATCTCTGATGTTAATATCTCATGCTCTCCACAGAGTGTAGGGACCTTTGGG 420
QY 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
DB 421 TATTGGCTCCAGAAATATGCTTCTCTGGAAAGCTTACAGACAAGTCAGACAGTATTCTTCC 480
```



of the cotton fiber  
Unpublished (2000)  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAATACGACTCACTATAGG  
High quality sequence stop: 690.

FEATURES  
    Location/Qualifiers  
        1..692  
            /organism="Gossypium arboreum"  
            /mol\_type="mRNA"  
            /strain="AXA"  
            /cultivar="8400"  
            /db\_xref="taxon:29729"  
            /clone="GA\_E40012C15f"  
            /tissue\_type="Fibers isolated from bolls harvested 7-10 dpa"  
            /lab\_host="E. coli"  
            /clone\_lib="Gossypium arboreum 7-10 dpa fiber library"  
            /note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
Alignment Scores:  
Pred. No.: 9.38e-25 Length: 692  
Score: 1047.00 Matches: 198  
Percent Similarity: 94.71% Conservative: 17  
Best Local Similarity: 87.22% Mismatches: 12  
Query Match: 30.32% Indels: 0  
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BG441204 (1-692)

Qy 247 LeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGlu 266  
Db 11 CTTCCGCTGGCTCACCTGGTATTCCTTGGTTCGAAAGACATTTAGCTATGAA 70

Qy 267 GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPhe 286  
Db 71 GAATAGCGAGACACGATGGCTTCGGAAGTTAACCTCTCTGGACAGGTGTTT 130

Qy 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306  
Db 131 GGGTACGTACACAAAGGATTCCTCCTAATGGGAAGAGTAGCAGTAAGCAACTCAAG 190

Qy 307 ValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgVal 326  
Db 191 GCTGGAAGTGGCAAGCGAGAGAGATTTCCAGGCTGAAGTTGAGATCATTAGCCGCTC 250

Qy 327 HisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeu 346  
Db 251 CATCAACAACATCTCTCTCATTTGGTCCGACTGTAATTCGGACAAATGAATGCTT 310

Qy 347 ValTyrGluPheValProAsnAsnLeuLeuGluHisLeuHisGlyGlyGlyArgPro 366  
Db 311 GTTTATGAGTTGTTCCAAACAACACCTTGGAGTTTCACCTGGCATGGGAAGGCGCATG 370

Qy 367 ThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyr 386  
Db 371 ACCATGAGTTGGCCCAAGGATGAATAATTCCTTTAGGATCTGCAAAAGGACTGGCATAT 430

Qy 387 LeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 406  
Db 431 CTTTCATGAAGATTGTCATCCTTAAGATCATTCCCGTGATATTAGCCCGCTAATATCTG 490

Qy 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426  
Db 491 TTGGAATTCAAAGTTTGAAGCAAGGTTCTGATTTTGGACTAGCGAAATTTGCTTCGAT 550

Qy 427 ThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyr 446

Db 551 GTCACACGACGCTCTCCACCGGTGATGGTACTTTCGGTATTATAGCCCTGAGTAT 610

Qy 447 AlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeu 466  
Db 611 GCTTCAAGTGGAAAGCTCACTATAATCAGATGTTTTCTCTCTCGGGGTGATGCTTTTG 670

Qy 467 GluLeuIleThrGlyArgArg 473  
Db 671 GAGTTGATTACCGGTACAGA 691

RESULT 10  
CD839231

LOCUS CD839231 723 bp mRNA linear EST 10-JUL-2003

DEFINITION RFO2.114104F010529 RFO2 Brassica napus cDNA clone RFO2114104, mRNA sequence.

ACCESSION CD839231

VERSION CD839231.1 GI:32521171

KEYWORDS EST.

SOURCE Brassica napus (rape)

ORGANISM Brassica napus

REFERENCE 1 (bases 1 to 723)  
Genoplaente, a major partnership french program in plant genomics  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

AUTHORS Genoplaente.

TITLE Genoplaente, a major partnership french program in plant genomics

JOURNAL Unpublished (2003)

COMMENT Contact: Genoplaente  
Genoplaente  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplaente' (<http://www.genoplaente.com>  
and <http://genoplaente-info.infobiogen.fr>).

FEATURES  
    Location/Qualifiers  
        1..723  
            /organism="Brassica napus"  
            /mol\_type="mRNA"  
            /cultivar="samourai (restored line)"  
            /db\_xref="taxon:3708"  
            /clone="RFO2114104"  
            /tissue\_type="anthers"  
            /clone\_lib="RFO2"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.37e-24 Length: 723  
Score: 1030.00 Matches: 199  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 29.83% Indels: 0  
DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x CD839231 (1-723)

Qy 449 SerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeu 468  
Db 2 AGCGAAAGCTCACCGAAGTCTGACGTTTCTCAITTTGGCGTTGTGCTTTGGAGCTC 61

Qy 469 IleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAsp 488  
Db 62 ATTACCAGGCGCTCACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 121

Qy 489 TrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAsp 508  
Db 122 TGGGACGACCATTTGCTTAACCGAGCATCTGAGCAAGGAGAGACTTTGAGGGTTTACGTAT 181

Qy 509 AlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAla 528  
Db 182 GCAAGATGAATAATGGGTATGACAGAGAGAGATGGCTGCGATGGTGTGCTGCTGCG 241

Qy 529 AlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeu 548  
 Db 242 GCTTGTTGCCATTCAGCTCGCGCAGAGACTCGCATGAGCCAGATTGTGCGTGGTTA 301  
 Qy 549 GluGlyAnValSerLeuSerAspLeuAnGluGlyMetArgProGlyGlnSerAsnVal 568  
 Db 302 GAAGGAATGTATCACTGTCTAGATCTTAACGAAGGGATGAGACCAGGTCAAGCAATGTA 361  
 Qy 569 TyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAnGluAspMetLys 588  
 Db 362 TACAGCTCATACGAGGAGCAGCCAGTATGACTCGAGCCAGTACAATGAAGACATGAAG 421  
 Qy 589 LysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsn 608  
 Db 422 AAGTTTAGAAATGGCTCTTTGGAACCTCAAGAGTACAACGCCACGGGTGAGTACAGTAAT 481  
 Qy 609 ProThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArg 628  
 Db 482 CCGACCAAGCGACTATGAGCTGTAGCCGTCTGGTTCAAGCAGCGAGGGCCAAACACACGC 541  
 Qy 629 GluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647  
 Db 542 GAAATGGAGATGGGGAAGATTAAAGAAACCGGTGAGGTTATAGTGGACCTTCTCTT 598

## RESULT 11

AY108241  
 LOCUS  
 DEFINITION Zea mays PC0134818 mRNA sequence.  
 ACCESSION AY108241  
 VERSION AY108241.1 GI:21211319  
 KEYWORDS HTC.

## SOURCE

ORGANISM

Zea mays

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

Hainey, C. F., Dolan, M., Miao, G. H., Vogel, J. M., Whitsitt, M. S., Arthur, L. W., Hanafey, M., Morgante, M. and Tingey, S. V.

TITLE

Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes

JOURNAL

REFERENCE

2 (bases 1 to 1053)

Coe, E. H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

COMMENT

If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

## FEATURES

source

1. 1053

/organism="Zea mays"

/mol\_type="mRNA"

/db\_xref="taxon:4577"

/clone\_lib="Maize Mapping Project/DuPont Consensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

## ORIGIN

Alignment Scores:

Pred. No.: 5.61e-24

Score: 1027.50

Length:

Matches: 1053

203

Percent Similarity: 79.93%  
 Best Local Similarity: 70.24%  
 Query Match: 29.76%  
 DB: 11  
 Gaps: 4  
 Conservative: 28  
 Mismatches: 46  
 Indels: 12

US-10-086-464-2 (1-647) x AY108241 (1-1053)

Qy 365 ArgProThrMetGluTyrSerThrArgLeuLysIle-AlaLeuGlySerAlaLysGlyLe 384  
 Db 6 CTTCCG-----GAGTGGCTCTAGATTAAAGATCATGTTGGGTGCTGCCAAGGGTTT 59  
 Qy 384 uSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspLysIleLysAlaSerAs 404  
 Db 60 AGCTTATCTTCATGAAGACTGCCATCCAAAGATCATCCATCGTGCATATAAGGCATCTAA 119  
 Qy 404 nIleLeuLysAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAl 424  
 Db 120 CATCTCTCTTGGACTTCCAAATTTGAAGCTAAAGTTGCTGATTTGGACATTCGAAAGTTT 179  
 Qy 424 aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaP 444  
 Db 180 TACTGATAACAACACCCATGTTTCGACAGAGTAAATCGGACACTTTGGGTATTTGGCACC 239  
 Qy 444 oGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValVal 464  
 Db 240 TGAGTATGCATCTTCTGGCAAGCTAACAGAAAAATCCGATGATTTTCTTCGCGAGTCAT 299  
 Qy 464 lLeuLeuGluLeuLysThrGlyArgArgProValAlaAsnValTyrValAspAs 484  
 Db 300 GCTTCTTGGACTTACTTCTGGCGGCGCAGCAGTGTGACCAACCAACCATATATATGATGA 359  
 Qy 484 pSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheG 504  
 Db 360 CAGCTTGGTGTGCTGGCAAGCCATTTACTGATCGGAGCACTTGAGATGGTGAATATGA 419  
 Qy 504 uGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetVal 524  
 Db 420 TGCTTTAGTCGATCCTCGGCTGGGAAAGGACTTCAATCTTAATGAGATGGCAAGAATGAT 479  
 Qy 524 lAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI 544  
 Db 480 AGCCTGTGCAGCTGCATGTGTACGCCATTCTGCACGCTCGCGCCAGCATATGAGTCAGGT 539  
 Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProG 564  
 Db 540 CGTTGGGCTTTGGAGGCAATGTCTTTGGAGGACCTTAATGAAGGTGTCGGCTGG 599  
 Qy 564 yGlnSerAsnValTyrSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAs 584  
 Db 600 CCATAGCCCGCTTCTTTGGGTCTATC---AGCAGCTCGATTACGATTTCTGCCAGTACAA 656  
 Qy 584 nGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrG 604  
 Db 657 CGAGGACATGAAGAAGTTCAAGAAGATGGCATTCACCAACAAC-----TATACACGAG 710  
 Qy 604 yGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGluG 624  
 Db 711 CCAATACAGCGCCCAACCAAGTGAATATGACAGATACCTGCTGTCATCAAGCAGCGAGG 770  
 Qy 624 yGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly----- 639  
 Db 771 CCACCAGACCGAAGAGATGGAGTCGGGTGCAATGAAGAAGGTGCTACAGTGGTGCTA 830  
 Qy 640 ----GlnGlyTyrSerGlyProSer 646  
 Db 831 CAGCTCAGGATACAGCGGAGCCTCG 855

## RESULT 12

AY108243  
 LOCUS  
 DEFINITION Zea mays PC0134814 mRNA sequence.  
 ACCESSION AY108243  
 VERSION AY108243.1 GI:21211321

AY108243 1016 bp mRNA linear HTC 16-OCT-2002

```

KEYWORDS      HTC.
SOURCE         Zea mays
ORGANISM       Zea mays
REFERENCE      1 (bases 1 to 1016)
AUTHORS        Hayney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
                Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE          Zea Mapping Project/DuPont Consensus Sequences for Design of
                Overgo Probes
JOURNAL        Unpublished (2002)
AUTHORS        Coe,E.H.
REFERENCE      2 (bases 1 to 1016)
TITLE          Direct Submission
JOURNAL        Submitted (25-APR-2002) Maize Mapping Project, University of
                Missouri, Columbia, MO 65211, USA
COMMENT        If you are interested in getting corresponding physical clones,
                these are publicly available from ZmDB and may be found by BLAST
                searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
                www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
                maize cDNA sequences is either Virginia Walbot, Stanford or Pat
                Schnable, Iowa State, then clones may be requested from ZmDB:
                www.zmdb.iastate.edu.
FEATURES       Location/Qualifiers
                1..1016
                /organism="Zea mays"
                /mol_type="mRNA"
                /db_xref="maizeDB:637888"
                /db_xref="taxon:4577"
                /clone_lib="Maize Mapping Project/DuPont Cornsensus
                Library"
                /note="this sequence is part of a project of EST
                assemblies resulting from the application of public
                contigs to seed DuPont contigs; this resource was
                assembled by DuPont as part of a collaboration for the
                overgo addressing of BACs in conjunction with the Maize
                Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.:      6,53e-24      Length:      1016
Score:          1025.00      Matches:    200
Percent Similarity: 82.14%      Conservative: 30
Best Local Similarity: 71.43%      Mismatches: 42
Query Match:    29.68%      Indels:     8
DB:             11           Gaps:       4

US-10-086-464-2 (1-647) x AY108243 (1-1016)
Qy      373 ArgLeuYsIlleAlaLeuGlySerAlaLysGlyLeuSerTyLeuHisGluAspCysAsn 392
Db      3 AGATTAAAGATCGCTCTGGTGCTGCCAAGGGTTAGCTTATCTTCATGAAGACTGCCAT 62
Qy      393 ProYsIlleIlleHisArgAspIleYsAlaSerAsnIlleLeuIleAspPheLysPheGlu 412
Db      63 CCAAAGATCATCCACCGCGCATTAAGGCATCTAACATTCCTCTGTGACTTCAAATTTGAA 122
Qy      413 AlalysValAlaAspPheGlyLeuAlalysIlleAlaSerAspThrAsnThrHisValSer 432
Db      123 GCTATGTTGCTGACTTTGGCTTGCAAGTTCTACTACTGATGATACACACCCATGTGTCA 182
Qy      433 ThrArgValMetGlyThrPheGlyTyLeuAlaProGluTyLeuAlaAlaSerGlyLysLeu 452
Db      183 ACAAGAGTAATGGGCACCTTTGGGTATTGGCACCAGTATGCAGCATCTGGCAAGCTC 242
Qy      453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 472
Db      243 ACAGAAAAATCCGATGTTATTTCTTTCCGAGTGTCATGCTTCTGTGACTTTACTGGGGCG 302
Qy      473 ArgProValAspAlaAsnValTyLeuValAspSerLeuValAspTyrAlaArgPro 492
Db      303 CGACCAATTGACACACACCAACCATATATGATGATGCAGCTTGGTTGACTGGCGCAAGGCCA 362

```

---

```

Qy      493 LeuLeuAsnArgAlaSerGluGlnGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsn 512
Db      363 TTACTGATGAGCGCTCGAGGATGGTGAATATATGATGCTTTGGTGGATCTCTCGCTCGGA 422
Qy      513 AsnGlyTyLeuAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAlaCysValArg 532
Db      423 AAGACTTCAATCTCTACGAGATGGCAAGATGATGATGATGATGATGATGATGATGATGATG 482
Qy      533 HisSerAlaArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552
Db      483 CATTCGCCACGTCGTCGCCACCAATGAGTCAGGCTCGTTCGGGCTCTGGAAGCGCAGCTG 542
Qy      553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyLeuSerSerTy 572
Db      543 TCTTTGGAGGACCTTAATGAAGGTGTTGCACTCGCCATAGCCCTCTTTGGGTCTCATAT 602
Qy      573 GlyGlySerThrAspTyLeuAspSerSerGlnTyAsnGluAspMetLysLysPheArgLys 592
Db      603 ---AGCAGCTCTGATTACGATTCGGGCGAGTACCAACGAGGACATGCAGAGTTTCAGGAG 659
Qy      593 MetAlaLeuGlyThrGlnGluTyAsnAlaThrGlyGluTyLeuTyLeuSerAsnProThrSer 612
Db      660 ACGGCGTTCAATAACAACACCTACACCGAGCAGC---CAATACAGCGCGCCCAACAGTGAA 716
Qy      613 TyGlyLeuTyProSerGlySerSerGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 631
Db      717 TACGGCCAGGTGCGCTGCTGATCAAGCGAGCGCGCCACGAGGAGGAGGAGGAGGAGGAG 776
Qy      632 MetGlyLysIlleLysArgThrGly-----GlnGlyTyLeuSerGlyProSer 646
Db      777 TCGGGTCGATGAAGAAGGTGGCTACAGCGGCTACAGCGGCTACAGCGGAGGAGGAGGAGCTCG 836

RESULT 13
BM358715
LOCUS     BM358715
DEFINITION  GA_Ea0012D16r Gossypium arboreum 7-10 dpa fiber library Gossypium
                arboreum cDNA clone GA_Ea0012D16r, mRNA sequence.
ACCESSION  BM358715
VERSION    1
KEYWORDS   EST.
SOURCE     Gossypium arboreum
ORGANISM   Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE  1 (bases 1 to 666)
AUTHORS    Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
                Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE       An integrated analysis of the genetics, development, and evolution
                of the cotton fiber
JOURNAL     Unpublished (2000)
COMMENT     Contact: Wing RA
                Clemson University Genomics Institute
                100 Jordan Hall, Clemson, SC 29634, USA
                Tel: 864 656 7288
                Fax: 864 656 4293
                Email: rwing@clemson.edu
                Total High Quality bases = 646
                Seq primer: TAATACGACTCATATAGGG
                High quality sequence stop: 666.
FEATURES    Location/Qualifiers
                1..666
                /organism="Gossypium arboreum"
                /mol_type="mRNA"
                /strain="AKA"
                /cultivar="8400"
                /db_xref="taxon:29729"
                /clone="GA_Ea0012D16r"
                /tissue_type="Fibers isolated from bolls harvested 7-10
                dpa"
                /lab_host="E. coli"

```

```

/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/notes="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 7,55e-24 Length: 666
Score: 1018.00 Matches: 191
Percent Similarity: 95.41% Conservative: 17
Best Local Similarity: 87.61% Mismatches: 10
Query Match: 29.48% Indels: 0
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BM358715 (1-666)

Qy 247 LeuProProSerProGlyLeuValValLeuGlyPheSerLysSerThrPheThrGlu 266
Db 11 CTTCCGGCTCGCTCACCTGGTATTTCTTAGTTTCTCGAAGACATTTTAGTATGAA 70
Qy 267 GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPhe 286
Db 71 GAATTAGCGAGACAACGGATGGCTCTCGGAAGTTAACTTCTTGGACAAGGTGGTTT 130
Qy 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306
Db 131 GGGTACGTACACAAGAGTTCTCCCTAATGGGAAGGAGTAGCAGTAAAGCAACTCAAG 190
Qy 307 ValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgVal 326
Db 191 GCTGGAAGTGGCAAGCGAGAGAAATTTTCAGCTGAAGTTGAGATCAATTAGCCGGTC 250
Qy 327 HisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeu 346
Db 251 CATCACAAACATCTCGTCTCAATTCGTCGGTACTGTAATTTCTGGGACAATAAGTCTT 310
Qy 347 ValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgPro 366
Db 311 GTTTATGAGTTGTGTCCAAACAACACCTTGGAGTTTCACTTGCATGGGAAGGGCGACTG 370
Qy 367 ThrMetGluTrpSerThrArgLeuValLeuAlaLeuGlySerAlaLysGlyLeuSerTyr 386
Db 371 ACCATGATTTGCGCAGCAAGGATGAAATTCCTTTAGATCTGCAAAAGGACTGGCATAT 430
Qy 387 LeuHisGluAspCysAsnProLysIleLeuHisArgAspIleLysAlaSerAsnIleLeu 405
Db 431 CTTTCATGAGATTGTCTCATCTCAAGATCATTCACCGTGATTAAGCGCGCTAATATCTG 490
Qy 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426
Db 491 TTGGAATTTCAAGTTTGAAGCAAGGTTGCTGATTTTGGACTAGCGAATAATTTGCTTCGGAT 550
Qy 427 ThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyr 446
Db 551 GTCAACACGCAAGCTCCACCGAGGTGATGGGTACTTTCGGGTATTTAGCCCTTGAGTAT 610
Qy 447 AlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValVal 464
Db 611 GCTTCAAGTGGAAAGCTCACTGATAAATCAGATGTTTCTTCCTCGGGTCATG 664

RESULT 14
CA072174 695 bp mRNA linear EST 23-SEP-2003
LOCUS SCCAM1004A11.9 AM1 Saccharum officinarum cDNA clone SCCAM1004A11
DEFINITION 5', mRNA sequence.
ACCESSION CA072174
VERSION CA072174.1 GI:34924325
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE 1 (bases 1 to 695)
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
```

TITLE  
JOURNAL  
COMMENT

The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: 004 row: A column: 11  
Seq primer: T7 Promoter Primer.

FEATURES  
source

1..695  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCCAM1004A11"  
/lab\_host="DH10B"  
/clone\_lib="AM1"

/notes="Organ: Apical meristem and tissues surrounding of  
mature plants; Vector: pSport1; Site 1: SalI; Site 2:  
NotI; An unidirectional cDNA library generated from  
[Apical meristem and tissues surrounding of mature  
plants]. cDNA was prepared from polyA+ mRNA using  
SuperScript Plasmid System Kit (Invitrogen). The  
double-strand cDNAs were fractionated in a separose  
CU-2B 40cm-columns and fragments sizing between 0.8 and  
1.5 Kb were directionally cloned into the vector. Details  
of each source of RNA and library construction can be  
obtained at http://sucest.lad.ic.unicamp.br/public"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,13e-23 Length: 695  
Score: 1013.00 Matches: 195  
Percent Similarity: 91.85% Conservative: 19  
Best Local Similarity: 83.69% Mismatches: 16  
Query Match: 29.34% Indels: 3  
DB: 13 Gaps: 1

US-10-086-464-2 (1-647) x CA072174 (1-695)

Qy 271 AlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHis 290  
Db 4 GCAACCAATGGGTTTCTGACGCTAATCTGCTGGGCAAGCGGTTTGGGTTTGTTCAC 63  
Qy 291 LysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGly 310  
Db 64 AAAGGAGTGTACCAATGGCAGAGGTTGCTTTAAGCAGTTAAGAGATGGAAGTGGC 123  
Qy 311 GlnGlyLysArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisArgHis 330  
Db 124 CAGGAGAACCGAGTTTCAGGCAGAGGTTGAGATTATCAGCAGAGTACATCAAGCAT 183  
Qy 331 LeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPhe 350  
Db 184 CTTGTATCTCTGTGCGCTATTGCAATTTCTGGAGCCAAACAGGTGCTGCTCAGAGTTT 243  
Qy 351 ValProAsnAsnLeuLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTyr 370  
Db 244 GTTCCAAACATACATACATGGAAATTCACATACATGAAAGGAGAGACCAACCTTGGATTGG 303  
Qy 371 SerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAsp 390  
Db 304 CCACAAAGACTAAATAATGCTCTGGGTTCTGCCAAGGATTTGGCATATCTTCATGAAGAT 363  
Qy 391 CysAsnProLysIleLeuHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLys 410  
Db 364 TGCCATCTTAAGATCATTCATCTCGTGATTAAGCCCTCAAAATATTTCTTCTTGAATCTAAGA 423

Qy 411 PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis 430  
 Db 424 TTTGAAGTAAAGTGGCAGATTTTGGACTTGCAAAATTCACCTTCTGTATACAAACACCCAT 483  
 Qy 431 ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly 450  
 Db 484 GTTTCACACAGAGTAATGGGCACATTTGGGTACCTAGCACCTGAGTATGCTGCTTCGGC 543  
 Qy 451 LysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThr 470  
 Db 544 AAGCTCACTGAGAAATCAGATGCTCTCTTTTGGAGTAATGCTTCTTTGAGCTAATAACT 603  
 Qy 471 GlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTyrAla 490  
 Db 604 GGGCGGCGTCTCT-----GGTAATTCACAGACAGCAGATGACAACTTGGGTGACTGGGCA 657  
 Qy 491 Arg-ProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502  
 Db 658 AGGGCTTTGATGAATAAAGCAATTTGAGGATGGTAAT 694

BU099573 715 bp mRNA linear EST 29-AUG-2002  
 WHE3309\_A09\_A1725 Chinese Spring wheat drought stressed root cDNA  
 library\_Triticum aestivum cDNA clone WHE3309\_A09\_A17, mRNA  
 sequence.

BU099573  
 BU099573.1 GI:22547372  
 Triticum aestivum (bread wheat)  
 Triticum aestivum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Poidea; Triticeae; Triticum.  
 1 (bases 1 to 715)  
 Anderson,O.D., Chao,S., Close,T.J., Crossman,C., Fenton,R.D.,  
 Lazo,G.R., Nguyen H.T., Pham,J., Rausch,C.J., Turuspekov,Y.,  
 Wilson,C., Woo,J. and Zhang,D.  
 The structure and function of the expressed portion of the wheat  
 genomes - Chinese Spring drought stressed root cDNA library  
 Unpublished (2002)  
 Contact: Olin Anderson  
 US Department of Agriculture, Agriculture Research Service, Pacific  
 West Area, Western Regional Research Center  
 800 Buchanan Street, Albany, CA 94710, USA  
 Tel: 5105595773  
 Fax: 5105595818  
 Email: oanders@nwp.usda.gov  
 Sequences have been trimmed to remove vector sequence and low  
 quality sequence with phred score less than 20  
 Seq primer: SK primer.  
 Location/Qualifiers  
 1..715  
 /organism="Triticum aestivum"  
 /mol\_type="mRNA"  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone="WHE3309\_A09\_A17"  
 /tissue\_type="root"  
 /dev\_stage="Full tillering stage"  
 /lab\_host="E. coli SOLR"  
 /clone\_lib="Chinese Spring wheat drought stressed root  
 cDNA library"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
 Site 1: EcoRI; Site 2: XhoI; Plants were grown under  
 normal conditions. Then drought stressed to 80%, 70% and  
 60% RWC at Texas Tech University (D. Zhang in Ht Nguyen  
 lab). Total RNA was prepared separately for roots  
 collected at the three different drought conditions. Equal  
 amounts of total RNA were pooled from all three samples,  
 poly(A) RNA were purified, one cDNA library was made, and  
 the cDNA clones were in vivo excised to give pBluescript  
 SK(-) phagemids in the TJ Close lab at the University of

California, Riverside (Fenton, Turuspekov). Plasmid DNA  
 preparations and DNA sequencing were performed in the OD  
 Anderson lab (all other authors)."

ORIGIN

Alignment Scores:  
 Pred. No.: 1.04e-22 Length: 715  
 Score: 983.00 Matches: 184  
 Percent Similarity: 89.70% Conservative: 25  
 Best Local Similarity: 78.97% Mismatches: 24  
 Query Match: 28.47% Indels: 0  
 Ds: 13 Gaps: 0

US-10-086-464-2 (1-647) x BU099573 (1-715)

Qy 276 SerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuPro 295  
 Db 11 TCCGATGCTAATCTTCTCGGGCAAGGTGGCTTTGGATTGTTTCAAAAGGAGTCTGCCA 70  
 Qy 296 SerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGlu 315  
 Db 71 GATGGCACAAGAGTTGCTGTGAAGCAATTAAGAGATGGAAGTGGGCGAGAGCGGTGAG 130  
 Qy 316 PheGlnAlaGluValGluIleIleSerArgValHisHisArgHisLeuValSerLeuVal 335  
 Db 131 TTCAGGCAGAGGTTGAGATTATCAGCCGAGTACATCATAAACATCTCGTGACATTTGGTT 190  
 Qy 336 GlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsn 355  
 Db 191 GGTATTATGCTTCTGAAGACAAGAGGTGCTTGTCTATGAGTTTGTTCCTCAATAACACG 250  
 Qy 356 LeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLys 375  
 Db 251 TTAGAATTCCATATACATGGAAGCGCGTGGACCACTATGAGTGGCTTCAAGACTACGT 310  
 Qy 376 IleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIle 395  
 Db 311 ATTGCTTTGGGTTCTCGAAGGGATTGGCGTATCTTCACGAAGACTGCCATCCAAAGATC 370  
 Qy 396 IleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysVal 415  
 Db 371 ATTCACTGTCACATAAAGGCATCAATATTTCTTCTGGATTACAGATGTGAAGCTAAGGTG 430  
 Qy 416 AlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVal 435  
 Db 431 GCAGATTTTGGACTTGCAAAAGTTAAACCTCTGATAATAACACTCATGTGTTCCACCAGAGTA 490  
 Qy 436 MetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLys 455  
 Db 491 ATGGGCACATTTGGGTACCTTGCCACAGATGATGCTTCTTCTGGCAAGCTNACTGAGNAA 550  
 Qy 456 SerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProVal 475  
 Db 551 TCAGATGCTCTTTCTTTTGGAGTAATGCTTCTCGAGTTAATAACTGGGCGCGTCTCTGTA 610  
 Qy 476 AspAlaAsnValTyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsn 495  
 Db 611 AGTTCAAAACAAGCGCATATGGATGACAGCTTGGTTGACTGGGCAAGCCCTTTGATGACA 670  
 Qy 496 ArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508  
 Db 671 CAAGCACTCGAGGATGGTAATCACGATGCTTTAGTGGAT 709

Search completed: April 25, 2004, 00:59:49  
 Job time : 3628 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: April 24, 2004, 20:30:41 ; Search time 133 Seconds  
(without alignment)  
2699.648 Million cell updates/sec

Title: US-10-086-464-2  
Perfect score: 3453  
Sequence: 1 MSSAPSGTGTGSPSPNST.....REMGKIKRTGCGYSGPSL 647

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Dgapop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10086464/runat\_23042004\_083016\_5642/app\_query.fasta\_1.839  
-DB=Issued Patents NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10086464@cgn1\_1\_85@runat\_23042004\_083016\_5642 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	756	21.9	3239	4	US-09-228-986-9
2	744.5	21.6	1593	4	US-09-602-472A-1
3	740	21.4	2868	4	US-09-228-986-4
4	712	20.6	1488	4	US-09-579-182-4
5	703.5	20.4	2513	4	US-09-228-986-13
6	678	19.6	2432	4	US-09-228-986-7
7	667.5	19.3	1559	4	US-09-602-472A-3
8	604.5	17.5	4104	3	US-08-881-706-1
9	602	17.4	3297	4	US-09-579-182-3
10	600.5	17.4	3094	4	US-09-228-986-11
11	581.5	16.8	3590	1	US-09-587-889-1
12	581.5	16.8	3590	4	US-09-016-434-1093

13	581.5	16.8	3590	5	PCT-US96-09193-1	Sequence 1, Appli
14	578	16.7	2114	4	US-09-602-472A-5	Sequence 5, Appli
15	575	16.7	2749	1	US-08-265-628-1	Sequence 1, Appli
16	572	16.6	2571	1	US-07-717-331F-9	Sequence 9, Appli
17	572	16.6	2833	1	US-07-717-331F-1	Sequence 1, Appli
18	567	16.4	1554	2	US-08-587-880A-24	Sequence 24, Appli
19	566.5	16.4	2389	4	US-09-228-986-1	Sequence 1, Appli
20	562	16.3	2336	4	US-09-228-986-10	Sequence 10, Appli
21	559.5	16.2	2638	4	US-09-228-986-8	Sequence 8, Appli
22	546	15.8	2749	1	US-07-717-331F-4	Sequence 4, Appli
23	545	15.8	966	1	US-08-447-185-2	Sequence 2, Appli
24	545	15.8	2443	1	US-08-447-185-3	Sequence 3, Appli
25	543	15.7	2686	4	US-09-228-986-3	Sequence 3, Appli
26	531.5	15.4	5733	2	US-08-473-553A-1	Sequence 1, Appli
27	486	14.1	2943	4	US-09-503-922-2	Sequence 2, Appli
28	460.5	13.3	2649	4	US-09-228-986-12	Sequence 12, Appli
29	430	12.5	2568	4	US-09-228-986-2	Sequence 2, Appli
30	401	11.6	1926	4	US-09-249-585A-2	Sequence 2, Appli
31	401	11.6	1926	4	US-09-410-399-3	Sequence 3, Appli
32	401	11.6	2580	3	US-09-050-863-2	Sequence 2, Appli
33	401	11.6	2580	4	US-09-359-081-2	Sequence 2, Appli
34	401	11.6	5452	2	US-09-130-114-1	Sequence 1, Appli
35	401	11.6	8705	4	US-09-647-344A-14	Sequence 14, Appli
36	401	11.6	9600	3	US-08-910-647-1	Sequence 1, Appli
37	401	11.6	9600	4	US-09-620-925-1	Sequence 1, Appli
38	401	11.6	10596	1	US-07-884-811-15	Sequence 15, Appli
39	401	11.6	10596	1	US-07-885-971-15	Sequence 15, Appli
40	401	11.6	10596	1	US-08-087-783A-15	Sequence 15, Appli
41	401	11.6	10596	1	US-08-194-088B-15	Sequence 15, Appli
42	401	11.6	10596	2	US-08-194-087-15	Sequence 15, Appli
43	401	11.6	10596	5	PCT-US93-04648-15	Sequence 15, Appli
44	401	11.6	16080	4	US-09-724-566A-48	Sequence 48, Appli
45	395.5	11.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-228-986-9  
; Sequence 9, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Nieuwenhuizen, Timothy  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9  
; LENGTH: 3239  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-09-228-986-9

Alignment Scores:	1.88e-23	Length:	3239
Pred. No.:	756.00	Matches:	190
Score:	51.93%	Conservative:	92
Percent Similarity:	34.99%	Mismatches:	132
Best Local Similarity:	21.89%	Indels:	130
Query Match:	4	Gaps:	17
DB:			
US-10-086-464-2 (1-647) x US-09-228-986-9 (1-3239)			
Qy	89	ProProalaProVal-ThrProThrArgAsnProProSerValProGlyProPr	108
Db	1759	CTCCCCAACCATTCGGCCCTATTACTTCATCCCTCCCTATGCTTTT-----CC	1811
Qy	108	oserAsnProSerArgGluGlySerProArgProProSerProProSe	128

```
Db 1812 AGACAAT-----GGAGGAACC-----1827
Qy 128 rProSerSerAspGlyLeuSerThrGlyValValGlyValGlyLeuAlaIleGlyGlyValAl 148
Db 1828 -----GCCATAAGCAAAAGGTGTGATAGTTGGATCGCTATTGGCGCACGGT 1874
Qy 148 aLeuLeuValIleValThrLeuIleCysLeuLeuCysLeuLysLysArgArgAspGly 168
Db 1875 TCTGGTCTTGGCCTTGTGTATTAGGGTTATATGCTATCGACAAAAGAAACGGGGGA 1934
Qy 168 uGluAspAlaTyrTrpValProPro-----ProProGlyProLysAl 184
Db 1935 GAAA---GCTCTGAGTTGACACACCTTCGCATCTCGGCACCCAGTGGGAAGATAG 1991
Qy 184 aglyGlyProTyrGlyGlyGlnGlnIntrpArgGlnGlnAsnAlaThrProProSe 204
Db 1992 CGGAGGAGCGCCACCACTGAAAGGAGCAGCATGG-----2025
Qy 204 rAspHisValValThrSerLeuProProProLysAlaProSerProProArgGlnPr 224
Db 2025 -----2025
Qy 224 oProProProProProPheMetSerSerSerGlyGlySerAspTyrSerAspAr 244
Db 2025 -----2025
Qy 244 gProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheTh 264
Db 2026 -----TTCTC 2030
Qy 264 rTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGl 284
Db 2031 CTATGATGAATTAAGAGTGTCACCAATAATTTCTCCGATAGCAATGAATTAGGCTTCGG 2090
Qy 284 yGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGl 304
Db 2091 AGGATACGGAAGGTGTACAGGGAGTCTTCTGTGATGCTATATATTAGCAATCAAAAG 2150
Qy 304 nLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLese 324
Db 2151 AGCTCAGCAGGGTTCGATCAGGGTGCACCGAGTTCAAGACAGAAATCGAGCTGCTTC 2210
Qy 324 rArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysAr 344
Db 2211 CGCGGTTTCATCACAAGAATCTGTGTGGCTCATAGGATCTGTTCGAGCAAGAGAGCA 2270
Qy 344 gLeuLeuValTyrGluPheValProAsnAsnAsnLeuLeuHisLeuHisGlyGluGl 364
Db 2271 GATGTGGTCTATGATATATATGCTTAACGGAGCGCTCAGGGATAGCTTGACAGAAATC 2330
Qy 364 yArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLe 384
Db 2331 AGGCATTATCTTGATGGAAGAGAGGCTTCGTATAGCTCTAGGTTCCGCTAGAGGACT 2390
Qy 384 uSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspLysAlaSerAs 404
Db 2391 AGCTTATTCGCAAACTCGCGAATCTCTCAATATTATCCACAGAGATGTCAAGTCCACCA 2450
Qy 404 nIleLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAl 424
Db 2451 TATCTTGTGGAGCAACATCTGACGGCCAAAGTTCGGATTTCCGTTTGTCCAACTGGT 2510
Qy 424 aSerAspThrAsnThr---HisValSerThrArgValMetGlyThrPheGlyTyrLeuAl 443
Db 2511 ATCGGACAGCGGAAGGGGCACGTTTCGACCAAGTGAAAGCGACGCTGGGCTATTGGA 2570
Qy 443 aProGlyTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVa 463
Db 2571 TCCCGAATACTACATGATGATCAACAGCTGACAGAAAGAGCGATGTGTACAGCTTCGGGT 2630
Qy 463 lValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAs 483
Db 2631 GGTATGCTTGAGCTCATCTGCAAGCAACCGATTGAGAGGGCAAGTATGTCGTCG 2690
```

```
Qy 483 pAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPh 503
Db 2691 CGAG-----ATTCGACCGCCATCGACAAGAACGACCAG-----GACTA 2729
Qy 503 eGluGly-----LeuAlaAspAlaLysMetAsnAsn---GlyTyrAspArgGluGl 519
Db 2730 CTACGGCGTGAGGAAATGATGGACCCGTCATCAGGAGCATGGGCTAC---CTCGTCGG 2786
Qy 519 uMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgPr 539
Db 2787 GTTCAGCAGGTTCTTGATTTGGCGATGTCGATGTCGAGGAGTCGCTCGGACCGCCC 2846
Qy 539 oArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGl 559
Db 2847 CACAATGAGCGAGTGTGAGGCGATCGAGCATCTGTTGCAG-----AACGA 2894
Qy 559 uGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAs 579
Db 2895 TGGATATACACACCAACTCGACGTCGCGATCCGATCGCTCGTCG-----GGACGGACTTGG 2945
Qy 579 pSerSerGln-----TyrAsnGluAspMetLysLysPheArgLysMe 593
Db 2946 GTCCAGCAGGGCGCTCTCGGCATCCGTACACAGATCCCTTACCAG-----2994
Qy 593 tAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTy 613
Db 2995 -----AAGGAAGTTAGCTATAGCATTCCTTTGATTATAGT-----GGTGGATA 3038

RESULT 2
US-09-602-472A-1
; Sequence 1, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; PRIORITY FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
US-09-602-472A-1

Alignment Scores:
Pred. No.: 2,9e-23 Length: 1593
Score: 744.50 Matches: 172
Percent Similarity: 57.07% Conservative: 66
Best Local Similarity: 41.25% Mismatches: 116
Query Match: 21.56% Indels: 63
DB: Gaps: 9

US-10-086-464-2 (1-647) x US-09-602-472A-1 (1-1593)
Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeu 254
Db 181 TCGCGGGTTCGAAAGACGGGTCCGACGGTCAGATT-----216
Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluLeuAlaArgAlaThrAsnGly 274
Db 217 -----GCTGCTCATACATTCCATTTCCGGGAGCTTGCAGCTGCAACAAACAAT 264
```





```
Qy 476 AspAlaAsnAsnValTyrValAspSerLeuValAspTyrAlaAspProLeuLeuAsn 495
Db 808 GATAATGCTCGAGCACCAGGAGCACAACCTTTGTCGATGGCTAGGCGGTGTTTC--- 864
Qy 496 ArgAlaSerGluGlnGlyAspGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyr 515
Db 865 -----AAAGATCGTAGAAGTTTCCGAGATGGCGGATCCATCGCTCAAGGGCGGTAT 918
Qy 516 AspArgGluGluMetAlaAlaCysAlaAlaCysValArgHisSerAla 535
Db 919 CCAATGCGTGTCTATATCAAGCACTTCAGTTCGAGCAATGTGTTACAGGAACAAGCA 978
Qy 536 ArgArgProArgMetSerGlnLeuValArgAlaLeuGluGlyAsnValSerLeuSer 555
Db 979 GCGACAAACCACTGATTGGCGACGTGGTGACAGCTCTTAACATACTTAGCTTCGCAACG 1038
Qy 556 ---AspLeuAsn-----GluGlyMetArgProGlyGlnSerAsnValTyr--- 569
Db 1039 TTTGACCCAAACGCCAACAGCGGTCAAAACAGTAGAAGTGGGAGTGGGCCCACTTTATC 1098
Qy 570 -----SerSerTyrGlyGlySerThrAspTyrAsp----- 579
Db 1099 AGAACAAAGGATGATCGAGAGAGCTTGGAGATGGGAGTAGCTGGATGCTTCGCAGAG 1158
Qy 580 -----SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMet 593
Db 1159 ACTCGAGTGGTTAGGCTCACCAGCACTCACAAGAACTCTCCTGATTACAGAAAGAGG 1218
Qy 594 AlaLeuGlyThrGlnGlyTyrAsnAlaThrGlyGluTyrSerAsnProThrSer--- 611
Db 1219 GATATG---GTGAGGGAATCAATGCAGGATCAGAAAGTGGGAGCGAGACAGGAGCGGG 1275
Qy 612 -----AspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThrArg 628
Db 1276 TCAGGTAGAAATCGGGATTA-----ACCGATTGGAAGGCAAGAAATCACAG 1323
Qy 629 GluMetGluMetGlyLysLeuLysArgThrGlyGlnGly 641
Db 1324 AGAGGGACCCGGCGAGTGTGGGAGATCATCGAGAGGC 1362
```

## RESULT 5

```
US-09-228-986-13
; Sequence 13, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2513
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-13
```

```
Alignment Scores:
Pred. No.: 2,21e-21 Length: 2513
Score: 703.50 Matches: 178
Percent Similarity: 49.52% Conservative: 81
Best Local Similarity: 34.03% Mismatches: 179
Query Match: 20.37% Indels: 85
DB: 4 Gaps: 14
```

US-10-086-464-2 (1-647) x US-09-228-986-13 (1-2513)

```
Qy 84 ThrThrProGlySerProAlaProValThrProThrArgAsnProProSer 103
Db 795 ACTGGTCTGCTCCTCCTCCGCCCACTATATCTCCACCT----- 833
```

```
Qy 104 ValProGlyProProSerAsnProSerArgGluGlySerProArgProProSerSer 123
Db 834 -----CCACCTTCTTAC-----AATCGGCCACCCCAAGTCC 863
Qy 124 ProSer-----ProProSerProSerSerAspGly-----Leu 134
Db 864 TCAATCGTGGTCCCTTTCATCAGGTGGCTCAAGGCTGGGAACAGCAATAAGAAATCTCTG 923
Qy 135 SerThrGlyValValValGlyLysLeuAlaLysGlyValAlaLeuLeuValLeuThr 154
Db 924 AGTGGTGGTCCATAGTAGGGTATAATATTTCAGATTATTTTTCAGTGTGTGTGCTGCTATA 983
Qy 155 LeuLeuCysLeuLeuCysLysLysLysArgArgArgAspGluGluAspAlaTyrTyrVal 174
Db 984 TTAGGAGTTATTTTATATGACGTAAGTCTCTCTAGAGAGAGCAGGAT----- 1031
Qy 175 ProProProProProGlyProGlyProLysAlaGlyProGlyGlyGlnGlnGln 194
Db 1032 -----GAAAGAAAACTA 1043
Qy 195 TrpArgGlnGlnAsnAlaThrPro-----ProSerAspHisValValThrSerLeuPro 212
Db 1044 AGCAATCGTGTGTCTTTTCACCCCTCTATCTCCCTCGATGCTGAATATTATTGAAAGAGAGT 1103
Qy 213 ProProProLysAlaProSerProArgGlnProProProProProProProPhe 232
Db 1104 CCAGAGCAAAAGTCAGCTCATCACCTCTTGAATCGCTCTTAAGCTCCCTCTTCGAA 1163
Qy 233 MetSerSerSerGly-----GlySerAspTyrSerAspArgProValLeu 247
Db 1164 CGCAACAAGTCTACAGGGGACAAAGGCTTCGGAAGTATTTTTTCAAGTAAGAGACTAAA 1223
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrGluGlu 267
Db 1224 AACCCA-----ATATCAGCAACTGAATATTCTATTGCGACAG 1259
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Db 1260 CTGCAATGCGACACAAATAGTTTGTAGTCAAGATANTCTTATTCGGGGGGTCTCTTGA 1319
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 1320 CGAATCTACCGAGCAGAGTTCAGAGTGAAGAGATTTTGGCAGTGAAGAAATTTGACACT 1379
Qy 308 Gly-----SerGlyGlnGlyGluArgGluPheGlnAlaGluLeuLeuSerArg 325
Db 1380 TCTACGCTGTCCCTTACAAAGCCCTGGAAGCTTCTCGACGCGAGTATCTAATATATCGGC 1439
Qy 326 ValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeu 345
Db 1440 CTACATCATCTTAACATTACAGAACTAGTGGTTATTTCACAGAACACATGACATACCTT 1499
Qy 346 LeuValTyrGluPheValProAsnAsnAsnLeu-----GluLeuHisLeuHisGlyGlu 363
Db 1500 CTTGTGTATGAATATTTTCGACAAATGGATCACTCTATCAGCTATTGTCACATGGCAGATGAG 1559
Qy 364 GlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly 383
Db 1560 ACTACTAGAAATTTGCTTGGAAACATTCGTGTAAGATTCGCTGGGTTCAGCTCGAGTT 1619
Qy 384 LeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSer 403
Db 1620 TTAGAGTATTTCGATGAGATTTGCTCTCCATCTATTGTGCAATAAAAAATTCAGTCTGCT 1679
Qy 404 AsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle 423
Db 1680 AATATTTTGTGTGATGATGATTTCAACCCCTCGTCTGCCAGCTGTGGAATTTGCG----- 1733
Qy 424 AlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443
Db 1734 GCCCTCAATCCAAATCTCGAGCGTTCAGGTTTCAGGTTTCATTTGGGTACAGTGCT 1793
```

```
Qy 444 ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVal 463
Db 1794 CCTGAATATGTCATGTCAGGAATCTATACAATGAAGACCGATGTGTATAGTTTGGAGTG 1853
Qy 464 ValLeuLeuGluLeuLeuThrGlyArgArgProValAspAlaAsnAsnValTyrValAsp 483
Db 1854 GTATGCTTGAGCTTTGACAGCGCGGAGCCCTGGGATAGTTCAAGAACAGGTCCAGAA 1913
Qy 484 AspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPhe 503
Db 1914 CAGTCATTAGTAAGATGGGCCACCCTCAGCTTCAT-----GATATT 1955
Qy 504 GluGlyLeuAla-----AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
Db 1956 GATGCATTGGCAAGAGATGGTTCATCCAGCTTTGAAGGGAAGTTATCTCGCAAAATCTCTC 2015
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 2016 TCGCGCTTGGTGATATTATTGGCTCTGCAATCAGCTGACCAAGATTCGTCCTCCA 2075
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 2076 ATGCTGAAGTGGTCAAGCATTTGGTTCGTATGATGCAACGGCTAGCTCAATAAGAGG 2135
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlySerThrAspTyrAspSer 580
Db 2136 ATGACAGAGATGAA-----ACTGCAGACCAAGATCCT 2168
Qy 581 SerGlnTyr 583
Db 2169 GCAGATTAT 2177
```

## RESULT 6

```
US-09-228-986-7
; Sequence 7, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2432
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-7
```

```
Alignment Scores:
Pred. No.: 2,44e-20 Length: 2432
Score: 678.00 Matches: 147
Percent Similarity: 63.05% Conservative: 68
Best Local Similarity: 43.11% Mismatches: 104
Query Match: 19.64% Indels: 22
DB: 4 Gaps: 8
```

US-10-086-464-2 (1-647) x US-09-228-986-7 (1-2432)

```
Qy 254 LeuValLeuGlyPheSerLysSerThrPheThrTyrGluLeuAlaArgAlaThrAsn 273
Db 1046 ATTTGCTAGGCAATTTCAG---AGATTCTCAATCAAGAGATTACGGTGTTCACATAAT 1102
Qy 274 GlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisGlyVal 293
Db 1103 AATTATTAGTACCAAGATATTATTAGGATGAGGATATGGAATTGTCTATAAAGATTC 1162
Qy 294 LeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer---GlyGlnGly 312
Db 1163 CTACAGATGGCACTATAGTACCAATAAAGTTGAAGATGTTAATGTGGAGGAGA 1222
```

```
Qy 313 GluArgGluPheGlnAlaGluValGluIleLeuSerArgValHisHisArgHisLeuVal 332
Db 1223 GAAATTCAATTTCAACAGAGGTGGAATATGATAGCTTGGCTGTGTAGAACTATTA 1282
Qy 333 SerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValPro 352
Db 1283 CGATTGTATGATTTTGCACAACTCCAGAGAGGCTTCTGGTCTATCTCCATCATGCCA 1342
Qy 353 AsnAsnAsnLeuGlu-----LeuHisLeuHisGlyGluGlyArgProThrMetGluTyrP 370
Db 1343 AATGGAAGTGGCGCTCTTGTCTTAGAGATCATATTAATGAAAAGCTTGCCTCGACTGG 1402
Qy 371 SerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAsp 390
Db 1403 CCTACTCGCAAGCTATAGCCCTTGGAGCAGCTAGGGAGCTGTATATATTTGCATGAGCAA 1462
Qy 391 CysAsnProLysIleIleHisAspIleLysAlaSerAsnIleLeuIleAspPheLys 410
Db 1463 TGTGATCCCAAGATTATTCACCGGATGTGAACAGCAAAATATATTACTGGATGAATAT 1522
Qy 411 PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis 430
Db 1523 TTTCAAGCTGTTGTGGAGATTTTGGGTAGCAAAAGCTCTTGGATCACAGGGATTCTCAT 1582
Qy 431 ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly 450
Db 1583 GTGACTACTGTGTTCGAGGAGCGTAGGTACATATGCCCAAGATACCTTTTCACCGGGA 1642
Qy 451 LysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThr 470
Db 1643 CAATCTTCAGAGAAAACATGATATTGGCTTTGGGATATTACTTGGTAATATTCTTGGAAC 1702
Qy 471 GlyArgArgProValAsp-----AlaAsnAsnValTyrValAspAspSerLeu 486
Db 1703 GGACAAAGGCTTTAGATTTTGGCAGGCTGCAAAAGCAAAAAGTTGTA-----ATG 1753
Qy 487 ValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeu 506
Db 1754 CTGATTTGGGTAAAGAGCTT-----CATCAAGAGAGAAGAGTTGCACCTTCCTT 1801
Qy 507 AlaAspAlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCys 526
Db 1802 GCTGATAAAGATCTTAAGGGCAATTTGATAGATTGAGCTGAGAGAGATGTTTCAGGTT 1861
Qy 527 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArg 546
Db 1862 TCTTTGTATGCACCAATTTTCAGCCTGGACATCGTCCAAAATGTGTGATGTTTGTAGA 1921
Qy 547 AlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSer 566
Db 1922 ATGTTGGAAGGTGAT---GGATTGCAGAACCGTGGGAAACATTCGCAAAAATTTGAAACC 1978
Qy 567 AsnValTyrSer-----SerTyrGlyGlySerThrAspTyrAspSer 580
Db 1979 CCCCAGATACAGAGTAAGTACAGATACCCATACATATTCGAGTGTGGTGAAGAAGATTCT 2038
Qy 581 Ser 581
Db 2039 TCT 2041
```

## RESULT 7

```
US-09-602-472A-3
; Sequence 3, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; CURRENT FILING DATE: 2000-06-23
```

```
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1559
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)...(1319)
US-09-602-472A-3
Alignment Scores:
Pred. No.: 4,396-20 Length: 1559
Score: 667.50 Matches: 165
Percent Similarity: 57.50% Conservative: 65
Best Local Similarity: 41.25% Mismatches: 117
Query Match: 19.33% Indels: 57
DB: 4 Gaps: 9
US-10-086-464-2 (1-647) x US-09-602-472A-3 (1-1559)
Qy 219 SerProProArgGlnProProProProProProProProProMetSerSerSerGlyGly 238
Db 131 TCACCAAAAGTCGGCCATGAAGTCCACCAT-CCGCCCTT----- 171
Qy 239 SerAspTyrSerAspArgProValLeuProProProProSerProGlyLeuVal----- 255
Db 172 -----CCAATATTCCAGATTATCTTCGGGGTGTAGTAGGAGAAA 213
Qy 255 ----- 255
Db 214 CGAGAAACAATGTTAATAATGCTGCTAGCTTAAGGAGGAGTCATCAGGGCGCGGATG 273
Qy 256 LeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPhe 275
Db 274 CTCAA-ATCGCGCGCAACACCTTCACATTCGGCGAGCTTGCGCGCGCCACCAATAATTT 332
Qy 276 SerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu--- 294
Db 333 CAGCCTGATGCTCTTAAAGGAGGGGTGTTGTTGTGTATATAGAGGACGCTTCAG 392
Qy 295 ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyLysArg 314
Db 393 AGTAGTGTCAAGTGTAGCTGTTAAACAGCTGGATAGAAATGGACTCCAAGTTAACCGT 452
Qy 315 GluPheGlnAlaGluValGluLeuIleSerArgValHisLysArgHisLeuValSerLeu 334
Db 453 GAGTTCTTGTGAAGTCTTATGCTAGCCTTTTACATCATCCAAACTTGGTCAATTTG 512
Qy 335 ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn 354
Db 513 ATTGGCTACTCGCTGCGAGCGACACGACTCTTGTGTATGATGTTTATGGCCTGGGT 572
Qy 355 AsnLeuGluLeuHisLeuHisGly-----GluGlyArgProThrMetGluTyrSerThr 372
Db 573 TCCTTGGNAGATACCTTCATGATGTTCCGCTGCACAGAGAACCATTAGATTGG---ACA 629
Qy 373 -ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAs 392
Db 630 CAGGATGAAGATAGCGGCTGT-GCAGCAAGGGGTGGAGTTCTTCATGATAAAGCTAA 688
Qy 392 nProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheG 412
Db 689 CCCACCAAGTATTATCGGGACTTCAATATCATCAACAACTTTTGTGGTAGGAGATTTC 748
Qy 412 uAlaLysValAlaAspPheGlyLeuAlaLysIleAla---SerAspThrAsnThrHisVa 431
Db 749 ACCAAGCTGCCAGCTTCGGGCTCGCGAAACTGGGCCCCACTGGAGCAACAGTCTCATGT 808
Qy 431 lSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLy 451
Db ----- 451
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1559
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)...(1319)
US-09-602-472A-3
Alignment Scores:
Pred. No.: 4,396-20 Length: 1559
Score: 667.50 Matches: 165
Percent Similarity: 57.50% Conservative: 65
Best Local Similarity: 41.25% Mismatches: 117
Query Match: 19.33% Indels: 57
DB: 4 Gaps: 9
US-10-086-464-2 (1-647) x US-08-881-706-1 (1-4104)
Qy 107 ProPro-----SerAsnProSerArgGluGlyGlySerProArgProProSer 122
Db 2350 CCACCGGCTAAGTCTTGAACAATCCTGGTCTC---TGTGGTTATCTCTTCCCGGGTGT 2406
Qy 123 SerProSerProSerProSerSerAspGlyLeu-----SerThrGly 137
Db 2407 GATCCTTCA-----AATGCAGACGGTTATGCTCATCATCAGAGATCTCATGGA 2454
Qy 138 ValValGlyLysLeuAlaIleGlyGlyValAlaLeuValIleValThrLeuLysCys 157
Db 2455 AGGAGACGAGCGTCCCTTGTGCTAGTGTGGCGATGGGATTTGTTGTTCTCTTTTGTGTGT 2514
; Sequence 1, Application US/08881706
; Patent No. 6245969
; GENERAL INFORMATION:
; APPLICANT: Chong, Joane
; APPLICANT: Li, Jianming
; TITLE OF INVENTION: Receptor Kinase BIN1
; FILE REFERENCE: 07251/022001
; CURRENT APPLICATION NUMBER: US/08/881,706
; CURRENT FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4104
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97)..(3687)
US-08-881-706-1
Alignment Scores:
Pred. No.: 4,4e-17 Length: 4104
Score: 604.50 Matches: 161
Percent Similarity: 47.21% Conservative: 84
Best Local Similarity: 31.02% Mismatches: 148
Query Match: 17.51% Indels: 127
DB: 3 Gaps: 14
US-10-086-464-2 (1-647) x US-08-881-706-1 (1-4104)
Qy 107 ProPro-----SerAsnProSerArgGluGlyGlySerProArgProProSer 122
Db 2350 CCACCGGCTAAGTCTTGAACAATCCTGGTCTC---TGTGGTTATCTCTTCCCGGGTGT 2406
Qy 123 SerProSerProSerProSerSerAspGlyLeu-----SerThrGly 137
Db 2407 GATCCTTCA-----AATGCAGACGGTTATGCTCATCATCAGAGATCTCATGGA 2454
Qy 138 ValValGlyLysLeuAlaIleGlyGlyValAlaLeuValIleValThrLeuLysCys 157
Db 2455 AGGAGACGAGCGTCCCTTGTGCTAGTGTGGCGATGGGATTTGTTGTTCTCTTTTGTGTGT 2514
```

```
QY 158 Leu-----LeuCysLysLysLysLysArgArgArgGlu 168
DB 2515 ATATTGGGCGTCATCTTTGGTAGAGAGATGAGAGAGAGAGAGAGCGC 2574
QY 169 GluAspAlaTyrValProProProProProGlyProLysAlaGlyGlyProTyr 188
DB 2575 GAGTTGGAGATGATGCG-----GAAGGACATGGMAACTCTGGCGATAGA 2619
QY 189 GlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValVal 208
DB 2620 ACTGCTACACACACCAATTTGAAGCTGACTGGTGTGAAGAAGCCTTCAGTATCAATCTT 2679
QY 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProPro 228
DB 2680 GCTGCTTTCCGAGAGCCATTGCGG----- 2703
QY 229 ProProProPheMetSerSerSerGlySerAspTyrSerAspArgProValLeuPro 248
DB 2703 ----- 2703
QY 249 ProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeu 268
DB 2704 -----AAGCTCACGTTTCCGATCTT 2724
QY 269 AlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyr 288
DB 2725 CTTCAAGGCTACCAATGGTTTCCATAATGATGATCTGATGGTTCTGCTGGGTTGGAGAT 2784
QY 289 ValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGly 308
DB 2785 GTTTCAAAGCGATTTGAAGATGGAAGCGGCTGCTATCAAGAACTGATTCATGTT 2844
QY 309 SerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleSerArgValHisHis 328
DB 2845 AGCGGTCAAGGTGATAGAGAGTTTCATGCGGAGATGGAACCACTTGGCAAGATCAACAT 2904
QY 329 ArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyr 348
DB 2905 CGAAATCTTGTCCTCTTCTTGTTATTGCAAGTTGGAGAGCGCGCTCTTGTGTTAAT 2964
QY 349 GluPheValProAsnAsnLeuLeu-----LeuHisLeuHisGlyGluGlyArgPro 366
DB 2965 GAGGTATGAGATGGAAGTTTAGAAGATGTTTTCGAAGACCCCAAGAAAGTGGGGTG 3024
QY 367 ThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyr 386
DB 3025 AAACCTTAATTTGTCACACGCGGGAAGATTGCGATAGGATCAGCTAGAGGCTTCTTTC 3084
QY 387 LeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 406
DB 3085 CTTCCACCAACTGCAGTCCGATATCATCCACAGAGACATGAATCCAGTAATGTGTG 3144
QY 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426
DB 3145 CTTGATGAGAATTTCGAAGCTCGGGTTTCAGATTITGGCATGGCGAGCTGATGAGTGG 3204
QY 427 ThrAsnThrHisValSerThrArgValMet---GlyThrPheGlyTyrLeuAlaProGlu 445
DB 3205 ATGATACGCAATTTAAGCGTCAGTACATTAGCTGTGTACACCGGGTTACGTTCTCCAGAG 3264
QY 446 TyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeu 465
DB 3265 TATTACCAAAAGTTTCCAGGTGTTTCAACAAAGGAGACGTTTATAGTTACGGTGTGCTCTTA 3324
QY 466 LeuGluLeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSer 485
DB 3325 CTCGAGCTACTCACCGGTAAACCGCAACCGGATTCACCGGAT---TTTGGAGATTAACAC 3381
QY 486 LeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyGly 505
DB 3382 CTTGTTGGATGGTGAACAG----- 3402
```

```
QY 506 LeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGlu----- 518
DB 3403 CACGCAAAACTGCGGATTAGCGATGTGTTTACCCGAGGCTTATGAAGAGATCCAGCA 3462
QY 519 -----GluMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAla 535
DB 3463 TTAGAGATCGAATTTTACAACATTTAAAGTTGCGGTGCGTGTGGATGATCGGCGCT 3522
QY 536 ArgArgArgProArgMetSerGlnIleValargAlaLeuGluGlyAsnValSerLeu--- 554
DB 3523 TGGAGACGACGACATGTTACAACTCATGCCCATGTTTAA-GGAGATACAAGCGGCTC 3581
QY 555 -----SerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
DB 3582 AGGATGATTTCACAGTCAACGATCAGATCAATAGAGATGGAGG----- 3626
QY 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAsp 586
DB 3627 -----GTTTCAGTACAATAGAGAT 3644

RESULT 9
US-09-579-182-3
; Sequence 3, Application US/09579182
; Patent No. 6500628
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: MNI-161
; CURRENT APPLICATION NUMBER: US/09/579,182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1224
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-579-182-3

Alignment Scores:
Pred. No.: 1,81e-17 Length: 1224
Score: 602.00 Matches: 153
Percent Similarity: 48.41% Conservative: 60
Best Local Similarity: 34.77% Mismatches: 157
Query Match: 17.43% Indels: 70
DB: Gaps: 9

US-10-086-464-2 (1-647) x US-09-579-182-3 (1-1224)

QY 128 SerProSerSerAspGlyLeuSerThrGlyValValValValGlyIleAlaIleGlyGlyVal 147
DB 28 TCTCGGAACACTCAATGGTTTGGAGCTATGGGAGATAATAGTATGTTCTCTCCGCGATA 87
QY 148 AlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysLysArgArgAsp 167
DB 88 TTCGTTGATGTTTAGCTATATATCGTATGCTGCTTACTTTCAGAGAAAAACCTCTAGATCT 147
QY 168 GluGluAsp-----AlaTyrTyrValProProProProProProGlyProLys 183
DB 148 TCTTCTAATCTAATCCCTGTTAGTCCGAGATTCCTCTAGTGTTCCTGAAGAGATTAAA 207
QY 184 -----AlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnTrpArgGln 197
DB 208 GAGATTAGAGTCGACGAGGTTTCTTCAAGCAATGGTGGG----- 246
QY 198 GlnAsnAlaThrProProSerAspHisValValThrSerLeuProProProPro----- 215
DB 247 ---RATGATACCCCTCTTATTAGTGAGAAATTTGGCGATAAAGAACCCGAAAAAGGGATA 303
QY 216 LysAlaProSerProProArgGlnProProProProProProProProPheMetSerSer 235
DB 304 AAGCAGAGTCAAGAAAAATGGCGGCTAGTAGCGCGCTCAGGCTCGTTTAATCACTTGGAGAAA 363
```

```
Qy 236 SerGlySerAspTyrSerAspArgProValLeuProProProSerPro-----Gly 253
Db 364 AAAGCGGATCGAGCGATCTCTGCTAATCCTTTGACAGCTCCATCTCTCTTCTGCTGGT 423
Qy 254 LeuVal-----LeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAla 269
Db 424 CTTCTCTGAGTTTCTCACCTTGGATGGGA---CATTTGGTTCACTCTTAGAGATCTTCAG 480
Qy 270 ArgAlaThrAnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrVal 289
Db 481 ATGCTACTAATCAGTTTCAAGGATATATCATCGGTGATGGTGGATATGGAGTTGTT 540
Qy 290 HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
Db 541 TACCGCGTAACTTGTAAATGGTACTCTCTGCTGTGTTAAAGATTGCTCAACAATTTA 600
Qy 310 GlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisArg 329
Db 601 GGACAAGCTGATAAAGACTTCAGAGTTGAAGTTGAAGCTATAGGTCAGGTTTCGACATAAA 660
Qy 330 HisLeuValSerLeuValGlyTyrCysTleAlaGlyAlaLysArgLeuLeuValTyrGlu 349
Db 661 AACTTGGTCCCGCTTCTCGGATATGTATGGAAGAAGACGAGAG-----705
Qy 350 PheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGlu 369
Db 705 -----705
Qy 370 TrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGlu 389
Db 706 -----CTCGGCTACCTTCACGAG 723
Qy 390 AspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPhe 409
Db 724 GCGATTGACCAAGTGTGTCAGACACATTAAGTCTAGTAACATCTGATTTGATGATGAC 783
Qy 410 LysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThr 429
Db 784 AAATTCATCTTAAATTTCTGACTTTGGACTTCTTAACTACTTTGGTGTGTAAGAGT 843
Qy 430 HisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSer 449
Db 844 TTTATAACTACTAGATGTTATGGTACCTTCGGTTACGTAGCTCCAGAGTATGCCAATTC 903
Qy 450 GlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIle 469
Db 904 GGTCTTCTGAATGAGAAAGCGATGCTACAGCTTCGGGGTTGTACTCTTGGAAAGCTATA 963
Qy 470 ThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTyr 489
Db 964 ACTGATAGATATCCGTAGACTATGCTCGTCCACCACCGAGGTACATTTGGTGGAGTGG 1023
Qy 490 AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAla 509
Db 1024 CTGAAGATGATGGTC-----CAACAAAGACGATCAGAAGAAAGTGGTTGATCCA 1071
Qy 510 LysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAla 529
Db 1072 AACCTTGAACAAACCACTTACAAAGTCTTTGAAAAGAACACATATTGACTGCTTTCAGA 1131
Qy 530 CysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGlu 549
Db 1132 TGTGTTGATCCAAATGCTGAGAAAAGACCGAGGATGAGCCAGTTGACGATGCTTGA 1191
```

## RESULT 10

```
US-09-228-986-11
; Sequence 11, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
```

```
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 3097
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
; US-09-228-986-11

Alignment Scores:
Pred. No.: 4,95e-17 Length: 3097
Score: 600.50 Matches: 173
Percent Similarity: 46.08% Conservative: 80
Best Local Similarity: 31.51% Mismatches: 172
Query Match: 17.39% Indels: 126
DB: 4 Gaps: 13

US-10-086-464-2 (1-647) x US-09-228-986-11 (1-3097)
```

```
Qy 66 ProSerThrProGlySerProProPro-----LeuPro 76
Db 1523 CCTGCAAGTTCTGACGCTCCAAAACACAGCTTAGACGGAGTTGTTCCCGACGCTTGGG 1582
Qy 77 GlnProSerProAlaProThrThrProGlySerProProAlaProValThrPro--- 95
Db 1583 AGAGCTGAAAAGACCTTCACCTACTGGACTTGGAGATAAAGCTGTCTACAG-GTACCTTAC 1641
Qy 96 -----ProThrArgAsnProProProSerValProGlyProPro 108
Db 1642 CAGATTCTTTGAAACAGACAGAGCTTGGAGGTCAGAACCTCAGGAAACTTGTGCTTTCT 1701
Qy 109 SerAsnProSerArgGluGlyGlySerProArgProProSerSerProSerPro--- 126
Db 1702 TCTCCACAACCGGTGC-GGTGATGCATCATCTAGTCTTCAATTGAGGCACCCCAAGTT 1760.
Qy 127 -----ProSerProSerSerAspGlyLeuSerThrGlyValValValGlyIleAla 143
Db 1761 ACAATAGTTCCGAGAGAAACAAAGGGGACATATCGTTAGCCATATATCTCGGAGCA 1820
Qy 144 IleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLys 163
Db 1821 GTCGAGGAGTATCACTAGCTATTTTACTCATCCGCTCTTCGTATTTCATGTACAGAGG 1880
Qy 164 ArgArgArgAspGluAspAlaTyrTyrValProProProProProGlyProLys 183
Db 1881 AGAGGAAGAACTGAA-----1895
Qy 184 AlaGlyGlyProTyrGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProPro 203
Db 1895 -----1895
Qy 204 SerAspHisValValThrSerLeuProProProLysAlaProSerProProArgGln 223
Db 1895 -----1895
Qy 224 ProProProProProProPheMetSerSerSerGlySerAspTyrSerAsp 243
Db 1896 -----ATGTCATATACGAA 1910
Qy 244 ArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
Db 1911 AGGCGAGTCCGACAGCTGAGAAACTGGAAT-----GCAGTAAGATTTT 1955
Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGln 283
Db 1956 TCCTACAAGAGATCAAAACAGCTCAACAACATTTAAAGAA-----GTCAATGGTTCAT 2009
Qy 284 GlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 2010 GGAAGTTTTCGATCCGTTGATACCTGCGGAAACCTTCCAGTTGGAACACTAGTTGCTGTGAAA 2069
```

```
Qy 304 GlnLeuLysValGlySerGlyGlnGluArgGluPheGlnAlaGluValGluLeuIle 323
Db 2070 GTGGCGTTTGATAAAACCACTTGGTGCAGATTCTTTTCATAAATGAGGTTGCTCTCTTA 2129
Qy 324 SerArgValHisLeuHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
Db 2130 TCACAGATCCCGCATCAGACATCTGTCTAGTCTGGAGGATTTGTATGATCGCAGCGT 2189
Qy 344 ArgLeuLeuValTyrGluPheValProAsnAenLeuLeuGluHisLeuHisGlyGlu 363
Db 2190 CAGATTTTGTCTATGAATATCTACCGGGTGGATCACTGGCTGATCAACTGTATGGTCCA 2249
Qy 364 Gly-----ArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAla 381
Db 2250 AACAGTAGGAATCTCTCACTAGCTGGTTCGTAGACTCAAGATTGCTGTGATGCTGCA 2309
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLys 401
Db 2310 AAAGACTGGACTATCTACATAATGAAGCAATCTCGAATCATACACCGAGACATCAAG 2369
Qy 402 AlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 2370 TGCAGTATATATTCTGACAGGAGATGAATGCAAGACTTTGGGACTTTGGGCTCTCT 2429
Qy 422 Lys-----IleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 2430 AAGCAAAATGATCCAGCCAGCGCA---ACTCACGTGACCACTGCTGCTCAAGGGCAGCACT 2486
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 2487 GGTACCTCCGACCTTGATATTTACTCCACCACCACTTACAGAAAGGACGCTCTAT 2546
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 2547 AGCTTTGAGTTGTCTTTGGAGCTCATCTGTGGACGAGAGCCGTTAAATCATTCAGGA 2606
Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 2607 ACTCAGATTCTCTCAATTGGTTTATGGCAAGCCCTACTTG-----CAG 2654
Qy 500 GlnGlyAspPheGluGlyLeuAlaAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db 2655 GCAGTGCATTGAG---ATAGTGATGAGAGTTAGGGGAGTTTCATGTGGAAAGC 2711
Qy 520 MetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgPro 539
Db 2712 ATGAGGAAGTGGCAAAATCGCTGTGAGTCTGTAGAGAGGATGATCACTAAGGCCA 2771
Qy 540 ArgMetSerGlnIleValArgAlaLeuGluGly-----AsnValSerLeuSerAspLeu 557
Db 2772 ACCATTGCACAGATACTGCTGTGCTCAAGAGCGCTTACAGCATTCAGCTCTCT----- 2825
Qy 558 AsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAsp 577
Db 2826 -----TATCTTGAGCCTCTGGACATGTGAAC 2852
Qy 578 TyrAspSerSerGlnTyrAsnGluAsp 586
Db 2853 TGAATCTACTGTATTACAAACCAAGAC 2879
```

## RESULT 11

```
US-08-587-889-1
; Sequence 1, Application US/08587889
; Patent No. 5654397
; GENERAL INFORMATION:
; APPLICANT: CAO, Zhaodan
; APPLICANT: CROSTON, Glenn E.
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
```

```
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,889
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-60916
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-587-889-1
```

## Alignment Scores:

```
Pred. No.: 3,48e-16 Length: 3590
Score: 581.50 Matches: 208
Percent Similarity: 38.68% Conservative: 86
Best Local Similarity: 27.37% Mismatches: 239
Query Match: 16.84% Indels: 229
DB: 1 Gaps: 28
```

US-10-086-464-2 (1-647) x US-08-587-889-1 (1-3590)

```
Qy 1 MetSerSerAlaProSerPro-GlyThrGlySerProProSerProProSerAsnSerTh 20
Db 80 ATGCCGGGGGGCGGGCCCGGGGAGCCCGCCAGCCCGGGCCGACATCTTCTGTAC 139
Qy 20 rThrThrThrProProAlaSerAla-----ProProPr 32
Db 140 GAGGTGCCCGCTGGGTCACTGTCCGCTTCTACAAAGTGTGGACGCCCTGGAGCCCGCC 199
Qy 32 oThrThrProSerSerPro----- 39
Db 200 GACTGGTGCCAGTTTCGCCCTGATCGTGGCCGCCAGCCGAGCTCGCGCTGTGCGAG 259
Qy 40 ----ProProSerThrIleProThrSerProProSerSerArgSerThrProSerAl 58
Db 260 CGCTCCGGGAGCGACGCGCCAGCGTCTGTGGCCCTGGATCAACCGCAACGCCCGTGTG 319
Qy 58 aProProProSerProPro-----ThrProSerTh 68
Db 320 GCCGACCTCGTGACATCTCTCACGCACCTGCAGCTGCTCCGTGGCGGGGACATCATCACA 379
Qy 68 rProGlySerProProProLeuProGlnProSerProProAlaProThrThrProGlySe 88
Db 380 GCCTGGCACCTCCCGCCCGCTTCGCTCCCGACGACCACTGCCCCGA-----GG 430
Qy 88 rProProAlaProValThrProThrArgAsnProProProSerValProGlyProPr 108
Db 431 CCCAGACATCTCTGCAACCCGCG-----AGGCCAGGGCTGGAGCCCCCGGA----- 479
Qy 108 oSerAsnProSerArgGluGlyGlySerProArgProProProSerProProProSe 128
Db 480 -----AGTTGCATCTCTCAGCTCCAGCTCCAGCTTCTCTC-TCCCCAGCTTT 519
```







Db 557 -----CTGGTTCCAGCCCTGCTTCCCTGTGGCTCCACCGCC 594  
Qy 228 oProProProMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuPr 248  
Db 595 ATCTCAGCCCTTCTTACCAAGCCAGGC-----CC 627  
Qy 248 oProProSerProGlyLeuValLeuGlyPheSerIysSerThrPheThrTyrGlu----- 266  
Db 628 AGAGAGCTCAGTGTCCCTCTCGAGGGAGCCGCCCTCTCCGTTTCTGGCCCCCTGTG 687  
Qy 267 -GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPh 286  
Db 688 TGAGATTCCCGGGCACCACCACTTCTCGAGAGCTCAAGATCGGGAGGGTGGCTT 747  
Qy 286 eGlyTyrValHisIstysGlyValLeuProSerGlyIysGluValAlaValIysGlnLeuLy 306  
Db 748 TGGGTGCTGTACCGGGCGGTGATG---AGAAACACGGGTGTATGCTGTGAAGAGGCTGAA 804  
Qy 306 sValGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGluI 322  
Db 805 GGAGAACCTGACCTGAGTGGAGTGCAGTGAAGCAGAGCTTCTCGACCGAGGTGGAGCA 864  
Qy 322 eIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAl 342  
Db 865 GCTGTCCAGGTTTCTGACCCCAACATTTGTGGACTTTGCTGGCTACTGTGCTCAGAACGG 924  
Qy 342 aLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGlu-----LeuHisLe 360  
Db 925 CTCTACTGCTGGTGTACGGCTTCTGCCCAACGGCTCCCTGGAGGACCGTCTCCACTG 984  
Qy 360 uHisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuIysIleAlaLeuGlySe 380  
Db 985 CCAGACCCAGCGCTGCCCACTCTCTCTGGCTCAGCGACTGGACATCTCTTCTGGGTAC 1044  
Qy 380 rAlaIysGlyLeuSerTyrLeuHisGluAspCysAsnProIysIleIleHisArgAspIl 400  
Db 1045 AGCCCGGCAATTGAGTTTCTACATCAGGAC---AGCCCCAGCTTCATCCATGGAGACAT 1101  
Qy 400 eLysAlaSerAsnIleLeuIleAspPheGlyPheGluAlaIysValAlaAspPheGlyLe 420  
Db 1102 CAAGAGTTTCCAAAGCTTCTGTGATGAGAGGTGACACCAAGCTGGGAGACTTTGGGCT 1161  
Qy 420 uAlaIysIle-----AlaSerAspThrAsnThrHisValSerThrArg-- 434  
Db 1162 GCGCGGTTACCGCTTTGCGGGTCCAGCCCCAGCCAGCAGCAGTGGTGGCCCGGAC 1221  
Qy 435 -----ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLe 452  
Db 1222 ACAGACATGCGGGCACCCTGGCTTACTTCCCGAGGAGTATACATCAAGACGGGAGGCT 1281  
Qy 452 uThrGlyIysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyAr 472  
Db 1282 GGCTGTGACACGGACACCTTTCAGCTTTGGGTGGTGTAGTGTAGAGACCTTGGCTGTGCA 1341  
Qy 472 sArgProVal-----AspAlaAsnAsnValTyrValAspAspSerLeuValAspTr 489  
Db 1342 GAGGCTGTGAAGACGACCGTGCAGACCAAGTATCTGAAGAC---CTGGTGGAA-- 1396  
Qy 489 pAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509  
Db 1397 -----GAGGAGGCTGAGAGGCTGGAGTGGCTTTTGAGAGACCCAGCAG 1440  
Qy 509 aLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAla----- 525  
Db 1441 CACACTGCAAGCAGGCTGTGGCTGCAGATGCTGGGCTGCTCCCATCGCCATCAGATCTA 1500  
Qy 525 ----- 525  
Db 1501 CAAGAAGCAGCTGGACCCCGCGGCCCTGCCCACTGAGCTGGGCTGGGCTGGG 1560  
Qy 526 ----CysAlaAlaAlaCysValArgHisSerAlaAlaArgArgProArgMetSerGlnIl 544  
Db 1561 CCAGCTGGCTGCTGCTGCTGCACCGCGGGGCCAAAGAGGCGCTCTCTATGACCCAGGT 1620

Qy 544 eValArgAlaLeuGlnGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564  
Db 1621 GTACGAGAGGCTAGAG-----AAGCTGCAGGCAGTGGTGGCGGGGTG---CCCGG 1668  
Qy 564 Y-----GlnSerAsnValTyrSerSer-- 571  
Db 1669 GCATTGGAGCCGCCAGCTGCATCCCCCTTCCCCGAGGAGAACTCTTACGTGTCCAG 1728  
Qy 572 -----TyrGlySerThrAspTyrAsp----- 579  
Db 1729 CACTGGCAGAGCCACAGTGGGGCTGTCTCCATGGCAGCCCTGGCAGCGCCATCAGGAGC 1788  
Qy 580 -SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGl 599  
Db 1789 CAGTGGCCAGCAGCAGCAGCAGCTGCAGAGAGGGCCCAACACGAGCCGCTGGAGAGTGACGA 1848  
Qy 599 uTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer-- 618  
Db 1849 G-----AGCTAGCGGCCTCTCTGTGCGCTCGGCTCTGGCACTTGACTTCCAGCTG 1902  
Qy 619 -----GlySerSerSerGluGlyGlnThrThrArgGl 629  
Db 1903 CCCTCTGGACCCAGCACCCCTCAGGGAGGCGCGCTGTCTCAGGGGACACGGC-AGGAG 1961  
Qy 629 uMetGluMetGly-----LysIleLysArgThrGlyGlnGlyTyrSerGlyPro 645  
Db 1962 AATCAGCTGGGGAGTGGCCAGGATCCCGGCCACAGCCGCTGGAAGGACTGGCCCT 2019  
RESULT 14  
US-09-602-472A-5  
; Sequence 5, Application US/09602472A  
; Patent No. 6608240  
; GENERAL INFORMATION:  
; APPLICANT: Bidney, Dennis L.  
; APPLICANT: Hu, Xu  
; APPLICANT: Li, Guihua  
; TITLE OF INVENTION: Sunflower Disease Resistance Genes  
; FILE REFERENCE: 35718/200630  
; CURRENT APPLICATION NUMBER: US/09/602,472A  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/140,876  
; PRIOR FILING DATE: 1999-06-24  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2114  
; TYPE: DNA  
; ORGANISM: Helianthus annuus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(2114)  
; OTHER INFORMATION: n = A,T,C or G  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1475  
; OTHER INFORMATION: n = A,T,C or G  
; NAME/KEY: misc\_feature  
; LOCATION: 1475  
; OTHER INFORMATION: n = A,T,C or G  
US-09-602-472A-5  
Alignment Scores:  
Pred. No.: 2,96e-16 Length: 2114  
Score: 578.00 Matches: 143  
Percent Similarity: 55.07% Conservative: 85  
Best Local Similarity: 34.54% Mismatches: 136  
Query Match: 16.74% Indels: 51  
Gaps: 15  
US-10-086-464-2 (1-647) x US-09-602-472A-5 (1-2114)



US-08-265-628-1

## Alignment Scores:

Pred. No.:	5,046-16	Length:	2749
Score:	575.00	Matches:	154
Percent Similarity:	49.80%	Conservative:	94
Best Local Similarity:	30.92%	Mismatches:	160
Query Match:	16.65%	Indels:	91
DB:	1	Gaps:	14

US-10-086-464-2 (1-647) x US-08-265-628-1 (1-2749)

Qy	130	SerSerAspGlyLeuSerThrGlyValValValGlyValLeuAlaLeuGlyValAlaLeu	149
Db	1327	AACGGAATGGGAAACCATATGCTGTTGTTGAGTTTGGTTT-----	1371
Qy	150	LeuValIleValThrLeuIleCysLeuLeuCysIleValLeuArgArgAspGlu	169
Db	1372	CTGCTTCTTATGATCATCTTCTGCG---CTCTGGAAGAGAAACAAACGAGCAAAACA	1428
Qy	170	AspAlaTyrTyrValProProProProGlyProLysAlaGlyProTyrGly	189
Db	1429	ACTGCAACATCTATTGTAAT-----	1449
Qy	190	GlyGlnGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValThr	209
Db	1450	-----CGACAGAGAAAC-----CAAGATTGCTAATGAAC	1479
Qy	210	SerLeuProProProProLysAlaProSerProProArgGlnProProProProPro	229
Db	1480	GGGATG-----ATACTATCAGCAAGACAGACAGTTCCT-----	1512
Qy	230	ProProPheMetSerSerGlySerAspTyrSerAspArgProValLeuProPro	249
Db	1513	-----ATAGAGAACAAACTGAGGAATTGGAACCTTCCATTGATA-----	1551
Qy	250	ProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAla	269
Db	1552	-----GAGTTGGAAGCTGTTC	1569
Qy	270	ArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrVal	289
Db	1570	AAAGCCACCGAAAATTTCTCCNAATGTAAACAACTCGGACAGAGGTTCGTTATGTT	1629
Qy	290	HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer	309
Db	1630	TACAAGGGTAGATTACTTGTGGGCAAGAAATTCGGGTAAAGGCTATCAAAAACGTCG	1689
Qy	310	GlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHisArg	329
Db	1690	GTTCAGGGAGCTGGTGAGTTTATGAATGAGGTGAGATTGATCGCAGGCTTCAGCATATA	1749
Qy	330	HisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGlu	349
Db	1750	AACTTGTCCGAAATTTCTGGCTGTTCATTTAGGCAGAGCAAGATGCTGTATATGAG	1809
Qy	350	PheValProAsnAsnLeuGluLeuHisLysGlyGluIleValGlyArg---ProThrMet	368
Db	1810	TATTAGAAAATTTAAGCTCGATCTTATCTCTTCGGAATTAACGAGCTCTACGTTA	1869
Qy	369	GluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHis	388
Db	1870	AATTGGAAGGACAGATTCAACATTACCAATGTTGTTGCTCGAGGACTTTTATATCTTCAT	1929
Qy	389	GluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAsp	408
Db	1930	CAAGACTCAGCGTTTAGGATATCCACAGATATGAAGTAACTAATTTTGTTCAT	1989
Qy	409	PheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsn	428
Db	1990	AAAAATATGACACCAAGATCTCGGATTTGGATGGCCAGAAATCTTTGCAAGGCGAG	2049
Qy	429	ThrHisValSerThrArg---ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla	447

Db	2050	ACTGAAGCTAACACAAGGAAGTGGTCGGAACCTTACGGCTACATGCTCCCGAGTAGCCCA	2109
Qy	448	AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu	467
Db	2110	ATGGATGGGTATTCTCGGAAAAATCAGATGTTTTTTCAGTTTTGGAGTCATTGTTCTTCAA	2169
Qy	468	LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal	487
Db	2170	ATTGTTAGTGGAAAAAGGACAGAGGATTCTTCAACTTTGAACCCAGCAAAACAATCTCTTA	2229
Qy	488	Asp-----TrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp-----	502
Db	2230	AGCTATGATGAGTCTACTGGCAGGAGGAGCGGTAGAAAATTTGTATCCAGTCATC	2289
Qy	503	PheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArg	522
Db	2290	GTAGATTCTATTGTCATCTATTACCAGCAAC-----TTTCAACCAAAAAGATTCTAAAA	2343
Qy	523	MetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgArgProArgMetSer	542
Db	2344	TGCATACAAATTGGTCTCTTGTGTTCAGAACGTCGACAGCATAGACCAACGATGTCG	2403
Qy	543	GlnIleValArgAlaLeuGluGlyAsnValSerLeuSerLeuAsnGluGlyMetArg	562
Db	2404	TCCGTGGTTGGATGCTTGGAGTGAAGCA-----ACGAGATTCCTGAGCCTACACCG	2457
Qy	563	ProGlyGlnSerAsnValTyrSerTyrGlyGlySer-----	575
Db	2458	CCAGTTTATCCCTCGGAAGAAGTCCCTTATGAAAAATAATCCTTCATCAAGTAGACATTGC	2517
Qy	576	-----ThrAspTyrAspSerSerGln	582
Db	2518	GACGACGACGAATCCTGGACGCTGAACCCAGTACACCTGCTCAGACATCGATCCCGGTAG	2577
Qy	583	TyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr	600
Db	2578	TACGA-AATCCGTTGAGAAAAGTTTCAGATAATTAACTATTGGGGTGACCCGGATAT	2630

Search completed: April 25, 2004, 01:02:38

Job time : 181 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 25, 2004, 00:01:06 ; Search time 638 Seconds  
(without alignments)  
4572.228 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSGTGPSPSPNST.....REMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2907579 seqs, 2254313464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPFO\_spool\_p/US10086464/runat\_23042004\_083017\_5726/app\_query.fasta\_1.839.  
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCUI=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DICALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10086464 @CGN 1 1 333 @runat 23042004 083017\_5726  
-NCPU=6 -ICPU=3 -NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARGTIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

RESULT 1  
US-10-086-464-1  
; Sequence 1, Application US/10086464  
; Publication No. US20020199218A1  
; GENERAL INFORMATION:  
; APPLICANT: GORING, Daphne R. et al.  
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES  
; FILE REFERENCE: P 25,762-A USA  
; CURRENT APPLICATION NUMBER: US/10/086,464  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 10/069,304  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: PCT/CA00/00966  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/149,466  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: US 60/159,122  
; PRIOR FILING DATE: 1999-10-13  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1944  
; TYPE: DNA

ALIGNMENTS





Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgProArg 540  
Db 1657 GCTCGCATGGTGTGCTGCGCGCTTGTGTGCGCCATTGAGCTCGCGCAGACCTCGC 1716  
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560  
Db 1717 ATGACCCAGATTGCGTGGCTAGAGGAATATGATCATCTGTGATCTTTAAAGAGGG 1776  
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580  
Db 1777 ATGACCCAGGTCAAAAGCAATGTATACAGCTCATACGGAGGAGCACCGATTATGACTCG 1836  
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600  
Db 1837 AGCCAGTACAATGAAGACATGAAGAAGTTTAGGAAAATGGCACATGGAACATCAAGAGTAC 1896  
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620  
Db 1897 AACGCCACGGGTGAGTACAGTAATCCGACCACTGATGACTATGACTGTACCCCTCTGGTTCA 1956  
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysLysLysLysArgThrGlyGln 640  
Db 1957 AGCAGCGAGGCGCAAAACCAACGCGAATGGAGATGGGGAAGATTAAAGAGAACCGGTGAG 2016  
Qy 641 GlyTyrSerGlyProSerLeu 647  
Db 2017 GGTATAGTGACCTTCTCT 2037

## RESULT 3

US-10-424-599-95479  
; Sequence 95479, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 95479  
; LENGTH: 2451  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_57229C.1  
US-10-424-599-95479

Alignment Scores:  
Pred. No.: 9e-140 Length: 2451  
Score: 2285.50 Matches: 450  
Percent Similarity: 80.86% Conservative: 74  
Best Local Similarity: 89.44% Mismatches: 97  
Query Match: 66.19% Indels: 27  
DB: 13 Gaps: 16

US-10-086-464-2 (1-647) x US-10-424-599-95479 (1-2451)

Qy 13 ProSerProSerAsnSerThrThrThrProProProAla-----SerAlaPro 30  
Db 246 CCCCTCCCTCGCCCT---GCAACTCCGTCGTCCACCTCCGCGGCAACTCCGTCGTCTCT 302  
Qy 31 ProProThrThrProSerSerProProProSerThrThrThrProSerProProPro 50  
Db 303 CCGCGCGGACACCTTCTTCCGACCTCCG-----TCAACTCTCTCTCCGCCCTCCG 356  
Qy 51 SerSerArgSerProSerAlaProProProSerProProProThrThrProGly 70  
Db 357 -----GGAGCTCCCTCCGCTTCTCCACCGTCCACTCTCTCTGCTTCTCTCCACCG 407

Qy 71 SerProProProLeuProGlnProSerProProAlaProThrThrThrProGlySerProPro 90  
Db 408 TCCACTCAACACCGCGCTCAACTTCGCGCCATCG---ACTTCTCGCGCTGCGCGCA 464  
Qy 91 -----AlaProValThrProThrArgAsnProProProSerValProGlyProPro 108  
Db 465 TCGCACTCGCGCGCTCGCTCCAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 524  
Qy 109 SerAsnProSerArgGluGlyGlySerProArgProProSerSerProSerProSerProSer 128  
Db 525 TCCCGGAGCTCGCCCTCTCTCCGTCGAGATCGAGACCGACCACTCGCTCGCTCCGCTCG 584  
Qy 129 ProSerSerAspGlyLeuSerThrGlyValValValGlyLysLysLysLysLysLysLys 148  
Db 585 TCTTCTCTCGTCGATATTCGACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 644  
Qy 149 LeuLeuValIleValThrLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 167  
Db 645 GTTCTTCTGT 704  
Qy 168 GluGluAspAlaTyrTyrValProProProProProProProProProProProProPro 186  
Db 705 GAAGAG-----TACTATGCTCCGCGCGCAACCGCGCGGCGGCGGCGGCGGCGGCGG 755  
Qy 187 ProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 206  
Db 756 GCATATGGT 812  
Qy 207 ValValThrSerLeuProPro-----ProProLysAlaProSerProProArgGln 223  
Db 813 GTGTGTCTCAATGATGCTCTCCAAAGCCATCGCCACCACTGCTCCACCGGCTTATGTGTGT 872  
Qy 224 ProProProProProProProProProPheMetSerSerSerGly-----GlySerAspTyr 241  
Db 873 CAACCTCCCGCCACCGCGCTCTCTTTCATCAGCAGCAGTGTGCGGATCTGGAATCAAACTAT 932  
Qy 242 SerAspArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysSer 261  
Db 933 TCAGCGGTGAATTTCT 992  
Qy 262 ThrPheThrTyrGluGluLeuAlaArgAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 281  
Db 993 ACATTCACTATGAGGAGTTGGCAGCGCAACTGATGCTCTCTGATGCCAACCTCTCT 1052  
Qy 282 GlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAla 301  
Db 1053 GGACAGGAGGATTTGGATATGTGCACAGAGGAATTTCTTCCCAACGGCAAGAGGTGGCA 1112  
Qy 302 ValLysGlnLeuLysValGlySerGlyGlnGluArgGluPheGlnAlaGluValGlu 321  
Db 1113 GTGAAGCAATTTGAAGGCTGGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1172  
Qy 322 IleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGly 341  
Db 1173 ATAATTAGCCGTGTCATCACACCATCTTGTCTTTCTTTGGTGGATATGATGATGATGATG 1232  
Qy 342 AlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuLeuLeuHisLeuHis 361  
Db 1233 TCCCAAGGCTGCTTGT 1292  
Qy 362 GlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysLysLysLysLysLysLys 381  
Db 1293 GGAAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1352  
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysLysLysLysLysLysLysLys 401  
Db 1353 AAGGAGCTGGCGTATCTTCATGAAGATTTGTCATCTTAAGATCATCTCATCTGATGATGAT 1412  
Qy 402 AlaSerAsnIleLeuIleAspPheGluAlaLysValAlaAlaAspPheGlyLeuAla 421  
Db 1413 GCTCCCAACATCCCTTCTGATTTTAAAGTTTGAAGCAAGGTTGAGATTTTGGTCTTGCA 1472  
Qy 422 LysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyr 441

```
Db 1473 AAGTTTTCTCTGATGTCAATACCCATGTTCTACTCGAGTGATGGGACATTTTGGGTAT 1532
Qy 442 LeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPhe 461
Db 1533 TTGGCTCAGAAATATGCTTCTAGTGGAAACTGACAGAAATCAGATGTTTCTCCTAT 1592
Qy 462 GlyValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 481
Db 1593 CGAGTCATGCTCTCGAGTTAATACCGACGACGCGCGCTCGATGATTAATAATCAAACTTTC 1652
Qy 482 ValAspSerLeuValAspTrpAlaAlaArgProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 501
Db 1653 ATCGAGGATAGTTTGGTAGACTGGCTAGGCTTGTCTCACACGAGCTTTTGGAAAGAGAT 1712
Qy 502 AspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAla 521
Db 1713 GATTTTGATTTCTATTATTGACCCAGGCTCCAGAAATGACTATGATCTTAATGAGATGGCA 1772
Qy 522 ArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgArgArgArgMet 541
Db 1773 CGATGTTGGCTTCTGCTCGGCTTGCAATTCGTATTCGGCAAGCGTCGACCAAGGATG 1832
Qy 542 SerGlnLeuValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMet 561
Db 1833 AGCCAGGTTGTTCCGCTCTCGAAGGAGATGCTCTCTAGCAGATCTTAACGAAGGAATT 1892
Qy 562 ArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSer 581
Db 1893 AGACCTGGACACAGCAGCTATGACAGTTCTCAT---GAAAGCTCAGATTAATGACACTGCA 1949
Qy 582 GlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsn 601
Db 1950 CAGTACAGGAAGACATGAAAAGTTTCAAAAATGGCATTTGGAACTCAGGAGTATGGT 2009
Qy 602 AlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSer 621
Db 2010 GCAAGCAGTGAGTACAGTCCGCTACAGTACAGTATGTTTAAACCCATCAGGCTCAAGT 2069
Qy 622 SerGluGly-----GlnThrThrArgGluMetGluMetGlyLysIleLysArgThr 638
Db 2070 AGTGAAGCAGACAGACCGCCAAACACACAGGGAATGGAATGAGAAAGATGAAG---AAC 2126
Qy 639 GlyGlnGlyTyrSerGlyProSer 646
Db 2127 AATCAGGTTTCAGTGGAAAGTTCT 2150
```

## RESULT 4

```
US-10-425-114-2326
; Sequence 2326, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Jiongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With.
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 2326
; LENGTH: 1724
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700209610_FLI
US-10-425-114-2326
Alignment Scores:
```

```
Pred. No.: 6.61e-113 Length: 1724
Score: 1870.00 Matches: 363
Percent Similarity: 79.85% Conservative: 53
Best Local Similarity: 69.67% Mismatches: 87
Query Match: 54.16% Indels: 18
DB: 13 Gaps: 9
US-10-086-464-2 (1-647) x US-10-425-114-2326 (1-1724)
```

```
Qy 139 ValValGlyLeuAlaIleGlyValAlaLeuLeuValIleValThrLeuLeuCysLeu 158
Db 1 GTCGTGGCGTGGCGGCTCGCGGCTCGTGTGCTGTGCTGCGCAGCTTCATCTGCGCTC 60
Qy 159 LeuCys-----LysLysLysArgArgAspGluGluAspAlaTyrTyrValProPro 176
Db 61 TGCTGTCTCCGCAAGAGCGCGCGCGCGCGCGCTCAGCACTACGATACCGATACCG 120
Qy 177 ProProProGlyProLysAlaGlyLysProTyrGlyGlyGlnGlnGlnTrpArg 196
Db 121 CCGCCCCCGCG---CCGTACAAGGAGGATCCATACGCTGGAACGTACCAAGATTGG--- 174
Qy 197 GlnGlnAsnAla-----ThrProProSerAspHisValValThrSerLeuPro---Pro 213
Db 175 CAGCAAAATCGCCTCTCTGCTCCACCCCTGAAACATGTGTCAAGATGCACCCCTCGCCT 234
Qy 214 ProProLysAlaProSerProArgGlnProProProProProProProProProPheMet 233
Db 235 CCGCACCATATGCCATCTCTCCACAGCGCGCGCGCGCCACCAACAGCATGTGTA 294
Qy 234 SerSerSerGly-----GlySerAspTyrSerAspArgProValLeuProProProSer 251
Db 295 AATAGTAGTGGTGTCTCTGCTCTCTGCTGCGAGATCTTACTCTGCTGCGAGATCTTACCTCCACCATCC 354
Qy 252 ProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAla 271
Db 355 CCGTGTCTGTCTCTGCTCTCTCGAAGACACATTCACCTACGAAAGATGTTTTCAGGGCT 414
Qy 272 ThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLys 291
Db 415 ACTGATGGATCTCGGATGCTATCTCTCTGGACAGGTGGTGGTGGTATGTTTCACAGA 474
Qy 292 GlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGln 311
Db 475 GGATTGCTGCTAATGCAAGAGATTGCTGTAACCACTGAAACTGGGAAGTGGCCAG 534
Qy 312 GlyGluArgGluPheGlnAlaGluValGluIleSerArgValHisHisArgHisLeu 331
Db 535 GCGCAGCGTGAGTTCCAGGCTGAGTTGAGATTATCAGCCGAGTACATCAAAACACCTT 594
Qy 332 ValSerLeuValGlyTyrCysIleAlaGlyValAlaLysArgLeuLeuValTyrGluPheVal 351
Db 595 GTGTCTTTGGTGGCTATTGCAATTTCTGGAGGCAAGAGGCTGCTTGTCTATGAGTTGTC 654
Qy 352 ProAsnAsnLeuLeuGluLeuHisLeuHisGlyGlyArgProThrMetGluTrpSer 371
Db 655 CCCAATAACACATTGGAAATTCCTTACATCGCAAGAGATCGAACCAACAATGGAGTGGCCT 714
Qy 372 ThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys 391
Db 715 GCTAGATTAAAGATCAGTTTGGGTGCTGCCAAGGGTTTAGCTTATCTTTCATGAAGACTGC 774
Qy 392 AsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPhe 411
Db 775 CATCAAAAGATCATCCATCGTACATTAAGCATCTTAACATCTTCTTGTGCTTCAATTT 834
Qy 412 GluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisVal 431
Db 835 GAAGCTAAGGTTGCTGATTTTGGACTTGGCAAGTTTCACTACTGATTAACAACACCATGTT 894
Qy 432 SerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLys 451
Db 895 TCGACAAGAGTAATGGGCACCTTTGGGTATTTGGCACCTGATGATGATCTTCTTGGCAAG 954
```



```
Qy 483 AspAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502
Db 2152 GATGACAGCTGGTTGACTGGCAAGGCCATTACTGATCGGAGCACTTGAGGATGGTGA 2211
Qy 503 PheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArg 522
Db 2212 TATGATGCTTTAGTGGATCCTCGGCTGGGAAGGACTTCAATCCCTAAATGAGATGGCAAGA 2271
Qy 523 MetValAlaCysAlaAlaCysValArgHisSerAlaArgAlaArgProArgMetSer 542
Db 2272 ATGATAGCTGTGACAGCTGCATGTGTACGCCATTCTGCACGTGCTGGCCACGATGAGT 2331
Qy 543 GlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArg 562
Db 2332 CAGGTGCTGGGCTTTGGAGGCCATGTCTTTGGAGGACCTTAATGAAGGTGTTGG 2391
Qy 563 ProGlyGlnSerAsnValTyrSerTyrGlyGlySerThrAspTyrAspSerSerGln 582
Db 2392 CCTGGCCATAGCCGCTCTTTGGGTCTATAC---AGCAGCTCCGATTACGATTCGGCCAG 2448
Qy 583 TyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAla 602
Db 2449 TACAACGAGCATGAAGAAGTTCAAGAAGATGCCATTCAACAACAAC-----TATACC 2502
Qy 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSer 622
Db 2503 AGCAGCCATACAGCGCCCAACGATGAATATGACAGATACCGTCTGCATCAAGCAGC 2562
Qy 623 GluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly----- 639
Db 2563 GAGGCGCACACAGACGAGAGATGAGTCGGTGCATGAAGAAAGGTGGCTACAGTGGT 2622
Qy 640 -----GlnGlyTyrSerGlyProSer 646
Db 2623 GGCTACAGCTCAGGATACAGCGGAGCCCTCG 2652

RESULT 6
US-10-425-114-33031
; Sequence 33031, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33031
; LENGTH: 2881
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLM017055B12_FLI
US-10-425-114-33031

Alignment Scores:
Pred. No.: 2,75e-110 Length: 2881
Score: 1833.00 Matches: 356
Percent Similarity: 79.80% Conservative: 51
Best Local Similarity: 69.80% Mismatches: 85
Query Match: 53.08% Indels: 18
DB: 13 Gaps: 9

US-10-086-464-2 (1-647) x US-10-425-114-33031 (1-2881)
Qy 150 LeuValIleValThrLeuIleCysLeuLeuCys-----LysLysLysArgArgAsp 167
```

```
Db 1139 GTGCTGCTCGCAGCTTCTATCTGCTCTGCTGCTCTCGCAAGAAGCGCGCGCGCGCG 1198
Qy 168 GluAluAspAlaTyrTyrValProProProProProProGlyProGlyProLysAlaGlyGlyPro 187
Db 1199 CCGCGCGCTCAGCAGTACCGATACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1255
Qy 188 TyrGlyGlyGlnGlnGlnGlnTyrArgGlnGlnAsnAla-----ThrProProSerAsp 205
Db 1256 TATGCTGGACGTACACAGAGTTGG---CAGCAAAATGGCGCTCTGCTCCACCCCTGAA 1312
Qy 206 HisValValThrSerLeuPro---ProProProLysAlaProSerProProArgGlnPro 224
Db 1313 CATGTGTCGAAGTACACCTTCGCTCGCCAGCATATGCCAATCGTCTCCACAGCGC 1372
Qy 225 ProProProProProProProPheMetSerSerSerGly-----GlySerAspTyrSer 242
Db 1373 CCAGCGCGCCACACACAGCGCATGCTAAATAGTAGTGTGGATCTGCTGTTCTTAATTA 1432
Qy 243 AspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThr 262
Db 1433 GTGGCGCAGATCTTACCTCCACCATCCCTGCTGCTGTTCTTGGCTTCTCGAAGACACA 1492
Qy 263 PheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGly 282
Db 1493 TTCACCTACGAAGAGCTGTTGAGGGCTACTGATGGATTCTCGGATGCTAATCTCTCTGA 1552
Qy 283 GlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaVal 302
Db 1553 CAAGTGTGTTTGGCTATGTTACAGAGGATGCTGCTTAATGCAAGAGATTTGCTGTA 1612
Qy 303 LysGlnLeuLysValGlySerGlyGlnGluArgGluPheGlnAlaGluValGluIle 322
Db 1613 AAACATTTGAACCTGGGAAGTGGCCAGGAGGCTGAGTTCAGGCTGAGTTGAGATT 1672
Qy 323 IleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAla 342
Db 1673 ATCAGCGGAGTACATCAACAACACCTTGTGCTTGTGGTATTGCTATTCTCTGGAGGC 1732
Qy 343 LysArgLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGly 362
Db 1733 AAGAGCTGCTTGTCTATGATGTTTGTCCCAATCAACATTTGGAATTTCCATTCATGCG 1792
Qy 363 GluGlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLys 382
Db 1793 AAAGTGCACCAACATGGAGTGGCTGTAGATTAAGATCAGTTTGGGTGCTGCCAG 1852
Qy 383 GlyLeuSerTyrLeuHisGluAspCysAsnProLysIleHisArgAspIleLysAla 402
Db 1853 GGTTTAGCTTATCTTCATGAAGACTGCCATCCAAAGATCATCATCGTGCATTAAGGCA 1912
Qy 403 SerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLys 422
Db 1913 TCTAACATTTCTTGTGACTTCCAAATTTGAAGCTAAGGTTGCTGACTTTGGACTTTGCA 1972
Qy 423 IleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeu 442
Db 1973 TTCCTACTGATTAACAACACCCATGTTCCACAAGAGTAATGGCACCTTTTGGGTATT 2032
Qy 443 AlaProGluTyrAlaAlaSerGlyLysLeuThrGlyLysSerAspValPheSerPheGly 462
Db 2033 GCACCTGAGTATGATCTCTTGGCAAGCTAACAGAAAATCCGATGTATTTCCTTCGGA 2092
Qy 463 ValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrVal 482
Db 2093 GTCATGCTTCTTGAGCTTATTACTGGCGCGCGAGCCAGTTCACACACCAACCAATATATG 2152
Qy 483 AspAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502
Db 2153 GATGACAGCTTGGTTGACTGGCAGGCCAATTACTGATGGAGCCTTGAGGATGGTGA 2212
Qy 503 PheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArg 522
```





```
Qy 255 lLeuGlyPheSerIysSerThrPheThrTyTGluGluLeuAlaArgAlaThrAsnGlyPh 275
Db 809 CCTCGGATTCAACAAGAGCACTTTTACTTCAACAAGAGCTTCGGCTGCAACAGAGGGTT 868

Qy 275 eSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyTrValHisGlyValLeuPr 295
Db 869 TACGGATCTAACCTTTTGGACAGAGGAGGATTTGGGTATCTCCATAAAGAGTCTTGCC 928

Qy 295 oSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgG1 315
Db 929 TAGCGGGAAGAGTAGGAGTTAAGAGTTTAAACCGGTAGCGGACAGAGAGAGGGA 988

Qy 315 uPheGlnAlaGluValGluLeuLeuSerArgValHisHisArgHisLeuValSerLeuVa 335
Db 989 GTTTCAAGCTGAGGTGCGATATCAATTAGCGGTGTGCATCATCGGTATCTGTTCTTGGT 1048

Qy 335 lGlyTyrcysileAlaGlyAlaLysArgLeuLeuValTyTrGluPheValProAsnAsnAs 355
Db 1049 TGGATATTGCGATAGCTGATGACAGAGAGATGTTGGTTTATGAGTTGTTCCTAAACAAAC 1108

Qy 355 nLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLy 375
Db 1109 TTTGGAATATCATCTTCATGGGAAAATCTTCGGTAATGAGTCTCCACTAGTTGCG 1168

Qy 375 sileAlaLeuGlySerAlaLysGlyLeuSerTyrlieuHisGluAspCysAsnProLysIl 395
Db 1169 TATCGCCTTAGGTGCTGCGAAAGGACTCGCTTACCTTCACGAAGACTGCCATCCTCGGAT 1228

Qy 395 eileHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysVa 415
Db 1229 CATTACCGCGACATCAAGTCTGAAAATATTCTCTTGGACTTCAACTTTGATGTATGGT 1288

Qy 415 lAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVa 435
Db 1289 GGCTGATTTGGATTAGTAGTAATCAATCTGTAACACACTCATGTATCTACTCGTGT 1348

Qy 435 lMetGlyThrPheGlyTyTrLeuAlaProGluTyTrAlaAlaSerGlyLysLeuThrGluLy 455
Db 1349 GATGGAACTTTCCGATATCTAGCTCCAGAAATATGCTTCAAGCGGTAAATTAACCGAGAA 1408

Qy 455 sSerAspValPheSerPheGlyValValLeuLeuGluLeuLeuThrGlyArgArgProVa 475
Db 1409 ATCCGATGTTTCTCTTACGAGGTATGTTATGGAATCTTAACTGGAACACACCGGT 1468

Qy 475 lAspAlaAsnAsnValTyTrValAspSerLeuValAspTrpAlaArgProLeuLeuAs 495
Db 1469 TGAT---AATAGCATCACCATGGACGACACCTTAGTAGATTGGGCTCGGCTCTTATGGC 1525

Qy 495 nArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTy 515
Db 1526 TCGCGCGCTAGAAATGGAACCTTTAATGAGCTGCGAGATGCGAGGCTTGAAGGCAACTA 1585

Qy 515 rAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAl 535
Db 1586 CAACCCGCAAGAAATGGCTCGAATGGTGACTTGTGCCCTGCTAGCATTCGTCTATCGGG 1645

Qy 535 aArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAenValSerLeuSe 555
Db 1646 CGGTAAAGCTCCAAAGATGACCGAGATAGTAAGAGCGTTAGAGGAGAGTCTCTTAGA 1705

Qy 555 rAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyTrSerSerTyTrGlyLys 575
Db 1706 TGCTTTAAACGAAGGTGTGAAGCCAGGACACAGATAACGTTTACGGGTCAATGGGAGCAAG 1765

Qy 575 rThrAspTyTrAspSerSerGlnTyTrAsnGluAspMetLysPheArgLysMetAlaLe 595
Db 1766 CTCGATTATTAGTCAGACATCTTCAATGCGAGACATGAAGAAATTCAGACAGATAGCTTT 1825

Qy 595 uGlyTyTrGlnGluTyTr-----AsnAlaThrGlyGluTyTrSerAsnProThrSerAs 612
Db 1826 CTCGAGCCCAAGAAATTCCTCCAGTCAGTGACTGTGGAAGGAACATCTAGTAAATGATCTAGAGA 1885
```

```
Qy 612 pTyrgly 614
Db 1886 TATGGGA 1892

RESULT 9
US-10-425-114-9117
; Sequence 9117, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9117
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700833950_FLI
US-10-425-114-9117

Alignment Scores:
Pred. No.: 2,26e-104 Length: 1424
Score: 1737.50 Matches: 331
Percent Similarity: 88.73% Conservative: 47
Best Local Similarity: 77.70% Mismatches: 41
Query Match: 50.32% Indels: 7
Gaps: 4
```

```
US-10-086-464-2 (1-647) x US-10-425-114-9117 (1-1424)

Qy 226 ProProProProProPheMetSerSerGly-----GlySerAspTy-SerAsp 243
Db 12 CCCCCACCAACCGCTCTTTCATCAGCAGCAGTCTGGATCAACAATTCAGGC 71

Qy 244 ArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263
Db 72 GGTGAATTTCTCTCTCTCTCTCCAGGAATTTTATTGGGGTTCTCTAAGAGCACATTC 131

Qy 264 ThrTyTrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuGlyGln 283
Db 132 ACGTATGAGGAGTTGGCAGCGCAACTGATGGCTTCTGTATGCCAACCTCTCTGGACAA 191

Qy 284 GlyGlyPheGlyTyTrValHisLysGlyValLeuProSerGlyLysGluValAlaValLys 303
Db 192 GGAGGATTTGATATGTGCACAGAGGAATTTCTCCCAACGGCAAGGAGTGGCAGTGAAG 251

Qy 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIle 323
Db 252 CAATTGAAGGCTGGAAGCGGCAAGCGGAGCGTGAATTTCCAAGCTGAAGTTGAGATATT 311

Qy 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyTrCysIleAlaGlyValLys 343
Db 312 AGCGGTGTCCATCAACAGCATCTTTGTTTCTTTGGTGTGATCTGATCATCTGATGGTCCAG 371

Qy 344 ArgLeuLeuValTyTrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGlu 363
Db 372 AGGCTGCTGTTTATGAATTTGTTTCCCAACAACACATTTGGAATTTCCATTTGATGAAAA 431

Qy 364 GlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly 383
Db 432 GGGAGACCTACCATGGGATTTGGCCCAAGCAAGATAAGAAATTCCTTTAGGATCTGCTAAGGGA 491

Qy 384 LeuSerTyTrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSer 403
```

Db 492 CTGGCGTATCTTCATGAAGATTGTTCATCTCAAGATCATCCATCGTGATATCAAAATCTGCC 551  
Qy 404 AsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle 423  
Db 552 AACATCCTTCGGATTATGTTTGAAGCAAGGTTGCAGATTTCGGTCTTGCAGAGTTT 611  
Qy 424 AlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443  
Db 612 TCTTCTGATGCAATACCACTACCTACTCGAGTGATGGGACATTTTGGGTATTTGGCT 671  
Qy 444 ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVal 463  
Db 672 CCAGAAATATGCTTCAGTGGAAATCTCACAGCAATCAGATGTTTCTCTCATGGAATC 731  
Qy 464 ValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAsp 483  
Db 732 ATGCTCCTTGAGTTAATACCGGACGCGGCGGAGTTGATAAATCAAACTTACATGGAG 791  
Qy 484 AspSerLeuValAspTrrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPhe 503  
Db 792 GATAGTTTGGTAGATTGGCTAGGCTTTTGCTCACACGAGCTTTGGAAGAGGATGATTTT 851  
Qy 504 GluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMetAlaArgMet 523  
Db 852 GATTCTATTTGACCCCAAGGCTCCAGAAATGACTATGATCTCATGAGATGGCAGCATG 911  
Qy 524 ValAlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGln 543  
Db 912 GTGGCTTCTGCGGCTTGATAGTCATGCTTCGCGCAAGCGTCGACCAAGGATGAGCCAG 971  
Qy 544 IleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgPro 563  
Db 972 GTTGTCCGCGCTCTCGAAGGAGATGCTCTCTAGCAGATCTGAATGAAGGAATAAGACCT 1031  
Qy 564 GlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyr 583  
Db 1032 GGACACACACTATATATGTTCTCAT---GAAAGCTCAGATTTATGACATGCACAGTAC 1088  
Qy 584 AsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThr 603  
Db 1089 AAGGAAGACATGAANAAGTTCAGGANAATGGATTGGGAATCTCAGGAGTATGGTGCAAGC 1148  
Qy 604 GlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGlu 623  
Db 1149 AGTGAGTACAGTCCGCTACAGTGAGTATGTTTAAACCCATCAGGCTCAAGTAGTGA 1208  
Qy 624 Gly-----GlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640  
Db 1209 GCACAGAGCCGCCAAACCCACAAAGGGAATGGAATGAGAAAGATGAAG---AACAATCAA 1265  
Qy 641 GlyTyrSerGlyProSer 646  
Db 1266 GGTTCAGTGGAAAGTTCT 1283

## RESULT 10

US-10-086-464-13  
; Sequence 13, Application US/10086464  
; Publication No. US20020199218A1  
; GENERAL INFORMATION:  
; APPLICANT: GORING, Daphne R. et al.  
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES  
; FILE REFERENCE: P 25,762-A USA  
; CURRENT APPLICATION NUMBER: US/10/086,464  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 10/069,304  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: PCT/CA00/00966  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/149,466  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: US 60/159,122  
; PRIOR FILING DATE: 1999-10-13  
; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 2025  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2025)  
US-10-086-464-13

Alignment Scores:  
Pred. No.: 6,08e-101 Length: 2025  
Score: 1687.00 Matches: 368  
Percent Similarity: 63.99% Conservative: 87  
Best Local Similarity: 51.76% Mismatches: 142  
Query Match: 48.86% Indels: 114  
DB: 14 Gaps: 25

US-10-086-464-2 (1-647) x US-10-086-464-13 (1-2025)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProSerProSerProSerProSerProSerThr 20  
Db 16 GTGATTCATCTCTGCCCCCTGAAACC---TCAATGGGACACCCCGTCAACGGAACA 72  
Qy 21 ThrThrThrProProAlaSerAlaProProThrThrProSerSerProProPro 40  
Db 73 TCG-----CGTCTAATGAGTCATCGCGGCCAACACCCACCTTCTTACACCA 123  
Qy 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro 60  
Db 124 TCATCA-----ATATCTGCTCTCCGCGCAGATATCTCCGCTCTCTTTCACCGCG 174  
Qy 61 ProSerProThr-----ProSerProGlySerProProPro 75  
Db 175 CTGCACCAACGCAAGAAACGTCACCTCTACATCTCCGCTCTCATCCGCCCTGTT 234  
Qy 76 -----ProGln-----ProSerProAlaPro-----Thr 84  
Db 235 GTAGCTAATCCGTCACCGCAGACTCCAGAGAACTCTTCTCCACCTGCACCTGAAGGCTCA 294  
Qy 85 ThrProGlySerProProAla---ProValThrPro-----ProThrArgAsn 99  
Db 295 ACTCCTTAACGCGCACCTGCACCAACCAACCGTCGAACCAATCACCAGGAAGACCA 354  
Qy 100 ProProPro-----Pro-----Pro-----Pro----- 102  
Db 355 ACTCCTCTCTCTCTGTCGCAATGATGACCGAACAAGAACCAATGCGGGGAACAACAC 414  
Qy 103 -----SerValProGlyProProSerAsnProSerArg-----GluGlyGly 116  
Db 415 AGAGACGCTCCACACCATCACACCGTCTGTCAGGGAACAGAACTTCCGGTGACGGTGC 474  
Qy 117 SerProArgProProSerProSerProProSerProSerProSerProSerProSerThr 136  
Db 475 TCACCTTCACACCTCGGTGATAAGCCCTCTCAGAAATAGTGGAGATTACAGACTCATCA 534  
Qy 137 GlyValValGlyIleAlaIleGlyGlyValAlaLeuValIleValThrLeuIle 156  
Db 535 TCG-----GGGCTT-----TTGCTTCTACTTGCAGTGTGTATT 567  
Qy 157 CysLeuLeuCysLysLysLysArgArgArgAspGluAspAla-----Tyr 172  
Db 568 TGCATCTGTTGCAACAGAGAGAGAGAGAAATCTCTCAGGTCAACCAACATGCACCTAC 627  
Qy 173 TyrValProProProProProGlyProLysAla---GlyGlyProTyrGlyGlyGln 191  
Db 628 TACAATAACAATCCTTATGGAGGAGCACCTCAGGTAATGTTGTTATTACAGGGA--- 684  
Qy 192 GlnGlnGlnTrpArgGlnGlnAsnAlaThrProSerAspHisValValThrSerLeu 211  
Db 685 -----ACACCTCAAGATCATGTGGTG---AATATG 711  
Qy 212 ProProProProLysAlaProSerProArgGlnProProProProProProPro 231

Db 712 GCTGTCAGGAGGTGGGAATTTGGGGTCCACAGCAACCTGTGTCTGGTCCT----- 762  
Qy 232 PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuProProSer 251  
Db 763 -----CACAGTGATGCTTCCAACTTAACCGGTGGAACGTGTATACCGTCACTCAA 813  
Qy 252 ProGlyLeuValLeuGlyPheSerIysSerThrPheThrTyrGluGluLeuAlaArgla 271  
Db 814 ---GCTGCAACTCTTGGTCAACAAAGCACTTTTCATCATGATGAAGTGTCCATTCGA 870  
Qy 272 ThrAsnGlyPheSerGluAlaAsnLeuGlyGlnGlyGlyPheGlyTyrValHisIlys 291  
Db 871 ACAGAAAGGTTTCCTCAGTCAGTCAAAATTTGTGAGCAAGGAGGATTTGGGTATGTTTCATA 930  
Qy 292 GlyValLeuProSerGlyGlyGluValAlaValIysGlnLeuLeuValGlySerGlyGln 311  
Db 931 GGAGTTCTGCTAGTGGCAAGAAAGTTGCAGTGAAGAGTCTTAAACTTTGGAAGTGGACA 990  
Qy 312 GlyGluArgGluPheGlnAlaGluValGluLeuIleIleSerArgValHisArgHisLeu 331  
Db 991 GGGGAACCGAGTTTCAAGACGAGGTTGATATCATTTAGCCGTGTCCATCGTCATCTC 1050  
Qy 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaIysArgLeuLeuValTyrGluPheVal 351  
Db 1051 GTTTCCTCTTGGTATATGATCTCTCGTGGTCAAGACCTTTGGTTATGAGTTTATA 1110  
Qy 352 ProAsnAsnLeuGluLeuHisLeuHisGlyGlyGlyArgProThrMetGluTrpSer 371  
Db 1111 CCTAACACACTCTTGAATTTTCATCTTCATGGAAGGTCGTCCGGTTTGGATGGCCT 1170  
Qy 372 ThrArgLeuIysIleAlaLeuGlySerAlaIysGlyLeuSerTyrLeuHisGluAspCys 391  
Db 1171 ACAAGATGAAGATTGCATTTGGATCGATGAGCCCTTGCATATTTGCATGAAGACTGT 1230  
Qy 392 -----AsnProIysIleIleHisArgAspIle 400  
Db 1231 AAGAAATCTTTATCTCATATATTTGTCATCAGTCACCTCGCATTTATCCAGAGATATC 1290  
Qy 401 LysAlaSerAsnIleLeuIleAspPheIysPheGluAlaIysValAlaAspPheGlyLeu 420  
Db 1291 AAAGCTGCACCAATCTTCTTCTGATTTTGTAGACCAAGGTGGCAGATTTTGGATTG 1350  
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440  
Db 1351 GCTAAGCTATCTCAAGACCACTATATCTCATGTCTCCACTCGCGTCATGGAACTTTTGA 1410  
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyIysLeuThrGluIysSerAspValPheSer 460  
Db 1411 TACTTAGCTCCAGAGTATGCATCAAGCGGAAGTTATCCGACAAATCTGTATGTTTCTCA 1470  
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480  
Db 1471 TTTGGAGTAATGCTTCTTGGAGCTATAACCGGAAGACCTCTCTGGATCTAACTGGA--- 1527  
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500  
Db 1528 GAAATGGAAGATAGCTTGGTAGATTGGCAAGGCGCTTTGTGTGTTGAAAGCAGCTCAAGAT 1587  
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaIysMetAsnGlyTyrAspArgGluGluMet 520  
Db 1588 GGAGATTACAACTTGGTGATCCAGTCTAGAGCTAAACTACAGTCATCAAGAGATG 1647  
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540  
Db 1648 GTTCAAATGGCTTCTTGTGACGTGCAGCAATCAGACATTCAGCAAGAAGACGGCTAAG 1707  
Qy 541 MetSerGln-----IleValArgAlaLeuGlu 549  
Db 1708 ATGACCCAGGTTCAAAAACTCATACCTTGTGTTGTTCTATTATTATTGACGACCTAGAA 1767  
Qy 550 GlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyr 569  
|||||

Db 1768 GGAGATATGTCAATGGATGATCTAAGTGAGGGAACAAGACCAGGACAAAGCACGTACTTG 1827  
Qy 570 SerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetIlys 589  
Db 1828 AGCCCGGAGCGGTGAGCTCAGAGTATGACCAAGCTCGTACACGGCAGACATGAAAAAG 1887  
Qy 590 PheArgIysMetAlaLeuGlyThrGlnGlyTyrAsnAlaThrGlyGluTyrSerAsnPro 609  
Db 1888 TTCAAGAACTCGCGCTAGAGAATAAGATATCAAGCAGT---GAATATGGTGA--- 1941  
Qy 610 ThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGlu 629  
Db 1942 ACAAGTCAGTATGCTTAAACCCCTTCTGCTTCAAGTAGTGAA-----GAA 1986  
Qy 630 MetGluMetGlyIysIleIysArgThrGlyGln 640  
Db 1987 ATGAATAGAGCTCAATGAACGCAATCCTCTCAG 2019  
RESULT 11  
US-10-086-464-12  
; Sequence 12, Application US/10086464  
; Publication No. US20020199218A1  
; GENERAL INFORMATION:  
; APPLICANT: GORING, Daphne R. et al.  
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES  
; FILE REFERENCE: P 25,762-A USA  
; CURRENT APPLICATION NUMBER: US/10/086,464  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 10/069,304  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: PCT/CA00/00966  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/149,466  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: US 60/159,122  
; PRIOR FILING DATE: 1999-10-13  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 2104  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-10-086-464-12  
Alignment Scores:  
Pred. No.: 6,31e-101 Length: 2104  
Score: 1687.00 Matches: 368  
Percent Similarity: 63.99% Conservative: 87  
Best Local Similarity: 51.76% Mismatches: 142  
Query Match: 48.86% Indels: 114  
Gaps: 25  
US-10-086-464-2 (1-647) x US-10-086-464-12 (1-2104)  
Qy 1 MetSerSerAlaProSerProGlyThrGlySerProSerProSerProSerAsnSerThr 20  
Db 52 GTGGATTCACTCTCTGCCCCGTAACCC---TCAATGGGACACCAACCGTCAACGGAACA 108  
Qy 21 ThrThrThrProProAlaSerAlaProProThrThrThrProSerSerProProPro 40  
Db 109 TCG-----CCGTCTAATGAGTCATCGCGGCAACACCAACCTTCTTCAACCAACCA 159  
Qy 41 ProSerThrIleProThrSerProSerSerArgSerThrProSerAlaProPro 60  
Db 160 TCATCA-----ATATCTGCTCTCCGCGCAGATATCTCCGTTCTTTTTCACCGCG 210  
Qy 61 ProSerProProThr-----ProSerThrProGlySerProProProleu 75  
Db 211 CTGCACCAACCAACGCAAGAAACGTCACCTCTTACATCTCCGTCCTCATCGCGCTGTT 270  
Qy 76 -----ProGln-----ProSerProProAlaPro-----Thr 84  
Db 271 GTAGCTAATCCGTCAACCGCAGACTCCAGAGAATCTTCTCCACCTGCACCTGAAGGCTCA 330



## US-10-424-599-57846

## Alignment Scores:

Pred. No.: 3,3e-91 Length: 2655  
Score: 1539.00 Matches: 341  
Percent Similarity: 55.19% Conservative: 95  
Best Local Similarity: 43.16% Mismatches: 141  
Query Match: 44.57% Indels: 213  
DB: 13 Gaps: 19

## US-10-086-464-2 (1-647) x US-10-424-599-57846 (1-2655)

```
Qy 2 SerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSer----- 19
Db 98 TCATCTTCCCAACA---AACACTTCCACCCACCCACCATCACCTCCAGTCTTCTCAGCCT 154
Qy 20 -----ThrThrThrProProProAlaSerAlaPro---ProProThrThr 34
Db 155 AATCAACAACAACAACAACAACAACCCCAATTCCTCGTCTCTCTCATCTCTCTCTGAC 214
Qy 35 ProSerSerProPro-----ProProSerThrThrProProSerProProSerSer 52
Db 215 CTTCTGCTCCACCTCTCTCTCTCTCTCTCTCAGGAGTGCATTAACCCCTCTCTCTCCATA 274
Qy 53 ArgSerThrPro-----SerAlaProProProSerProProThrProSerThrPro 69
Db 275 TCACGTCTCTCTCCACCATCTACACCAACCACTTTGCCCCCTCCATCACCACT 334
Qy 70 GlySerProProProLeuPro-----GlnPro 78
Db 335 GATCTCTCCACCACTTCCACCTGCATCTCCACCCCAACAGTTACAACTCTCCCCC 394
Qy 79 SerProProAlaProThrThrProGlySerProProAlaProValThrProProThrArg 98
Db 395 TCACCACCAAGTTACAACTCTCCCCCTCACCAACAGCTACCACTCCCTCTCCAGCT 454
Qy 99 AsnProProSerValPro-----GlyProProSerAsnProSerArgGluGly 116
Db 455 GAACTCCACCTTCCCTTCCAAATTTGTCTCCACCATCTCCACCAAGCC-----Ggt 505
Qy 117 SerProAlaPro-----ProSerSerPro 124
Db 506 TCCCTCTCTCTCAATCACCTCCCAACAATTCACCTCTCTCTCGACCGATTCGCCA 565
Qy 125 SerProProSerProSer----- 130
Db 566 TCTCTCTCTCTCCGCGCAATTTTCCAAGGCCCAACCAACCACTAGAACTCTCCAGAGAA 625
Qy 130 ----- 130
Db 626 GAAATCCACCAAAACTACTCTTCAATGATCTCTCCATCATCTCTCTGAAACTCTCT 685
Qy 130 ----- 130
Db 686 CCTAAACCTCTCTCTGATGTTCCCTCCATCCACATTCCTTCACTCTCTCTCTCA 745
Qy 130 ----- 130
Db 746 GACCTTCCAGGATCTTCACTCCAGCTTCTTTCCTGTGATCCCACTAATAAACAAGTG 805
Qy 130 ----- 130
Db 806 GTGGGGGTCCGAGGTGTGCTACCTCTCTTCCAACTGTGAACCCACTGTAGACCT 865
Qy 131 -----SerAspGlyLeu 134
Db 866 ACTAATGATGTACTTAACAGTATGCTTCAACAACAACAACCTTCCATCTCTGGGGTGT 925
Qy 135 SerThrGly-----ValValValGlyIleAlaIleGlyGlyValAlaLeuValIle 152
Db 926 AGCACTGGAGGATCTGTGGCTATTGGAATTTGTAGTTGTTTTATTGTCTCTCAGCCTCT 985
Qy 153 ValThrLeuIleCysLeuLeuCysIleValValValValValValValValValValVal 172
```

```
Db 986 GTTATGCTGTGCTGTTTGCACAGAGAAAGAGAGGAAACAGGATCAAGAGGTAGT 1045
Qy 173 TyrValProProProProProGlyProLysAlaGlyGlyProTyrGlyGlyGlnGln 192
Db 1046 TATGCTGCTCTCTCTCCA----- 1063
Qy 193 GlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuPro 212
Db 1064 -----TTTACCTCATCCCAATTCAGGTACCTTATTC 1096
Qy 213 ProProProLysAlaProSerProProArgGlnProProProProProProProPro 232
Db 1097 TTGAGGCGCAGTCTCCGGCC-----AACTTT 1123
Qy 233 MetSerSerSerGlySerAspTyrSerAspArgProValLeuProProProSerPro 252
Db 1124 TTAGGTAGTGTCTGCTGAGTGAATTT-----GTATATTCTCCATCAGAGCCT 1171
Qy 253 GlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaAlaThr 272
Db 1172 GGTGCTGA-----AGTAGTTCAAGATCATGTTTCATATGAAGAACTATTTCAGAGCTACA 1228
Qy 273 AsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisIleGly 292
Db 1229 AATGGTGTTCAGCACAAATTTGTTGGGAGAGGTGGATTTGCTGTGTATTATAGGT 1288
Qy 293 ValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGly 312
Db 1289 TTGCTGATAGATGAAGAGAGTAGCTGTGAAACAGCTCAAAATTTGGTGGTGGCAAGG 1348
Qy 313 GluArgGluPheGlnAlaGluValGluIleIleSerArgValHisIleAspHisIleVal 332
Db 1349 GAACGCCAATTCAGGCGCAGAGTTGAGATTATTAGCCGTGATCATCATCTGCTGTT 1408
Qy 333 SerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValPro 352
Db 1409 TCTTTAGTTGGTACTGTATATCCGACATCAGAGATTGCTTGTATATGACTATGTTCCC 1468
Qy 353 AsnAsnLeuLeuGluHisLeuHisGlyGlyArgProThrMetGluTrpSerThr 372
Db 1469 AACGATATCTTCTATTACCATCTCCAGGTGAAATAGACCACTTAGATTGGCTACC 1528
Qy 373 ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGlyAspCysAsn 392
Db 1529 AGAGTCAAGGTGCTGCTGTCAGCTCGTGGATAGCTTACTTGCATGAAGACTGTCTAT 1588
Qy 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412
Db 1589 CCAGCATATTATCATCAGATATTAGTCATCAAAACATCTCTACTTCTCAACTATGAA 1648
Qy 413 AlalysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer 432
Db 1649 GCTCAAGTTTCGGACTTTGGGCTTGCAAAATTTGGCAATTGTCATTAATAATACATGTA 1708
Qy 433 ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 452
Db 1709 ACAGTGTAAATGGGAACCTTTGGGTATATGGCACCAAGATATGGCAAGTGGAAACCTT 1768
Qy 453 ThrGlyLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 472
Db 1769 ACTGAAAGTCTGATGATATTTCTTTGGGGTGTGCTTTTGGAGCTAATTTACAGGTGG 1828
Qy 473 ArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAlaArgPro 492
Db 1829 AAGCTGTAGATGCTCTCAACCAATTTGGTATGAGAGCTTGGTGTGAATGGGCTCGACCT 1888
Qy 493 LeuLeuAsnArgAlaSerGlnGlnGlyAspPheGlyGlyLeuAlaAspAlaLysMetAsn 512
Db 1889 CTGTGTGAGGAGGACCTTTGCAATGAGACCTTTGAAATTTTGGTGGATCCAGACTGGGG 1948
Qy 513 AsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532
```

Db 1949 AAGAACTACGATAGAAATGTTTCGGATGATCGAGGTCGCGAGCCTGTGTACGC 2008  
Qy 533 HisSerAlaAatGArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552  
Db 2009 CACTCATCGTGAAGAGACCGCATGAGTACGGTGGTGAGAGCTTTAGATTCTCTGGAT 2068  
Qy 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr 572  
Db 2069 GAGTTTACGGATCTCAATAACGGAATGAACCGGACAGAGTTCGGTG----- 2116  
Qy 573 GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys 592  
Db 2117 -----TTTGAATCGCGCAGCAATCTGCACAAATCAGAATGTTTAGGAGG 2161  
Qy 593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp 612  
Db 2162 ATGGCTTTTGGAGGCAAGAT-----AGTCCCGGTTTCTCAATGAGTCTCAGAGTAGC 2215  
Qy 613 Tyr-GlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluMetGluMe 632  
Db 2216 TGGAGGAGTA-----GAGATCAGCACGCCCAACAACACTGTCTCTC 2251  
Qy 632 tGlyLysIleLysArgThrGlyGlnGly 641  
Db 2252 TCCCAAAATAAAA-----CTGGGCATGGGA 2276

## RESULT 13

US-10-425-114-7339  
; Sequence 7339, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53131)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 7339  
; LENGTH: 2655  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700652889\_FLI  
US-10-425-114-7339

Alignment Scores:  
Pred. No.: 3,3e-91 Length: 2655  
Score: 1539.00 Matches: 341  
Percent Similarity: 55.19% Conservative: 95  
Best Local Similarity: 43.16% Mismatches: 141  
Query Match: 44.57% Indels: 213  
DB: 13 Gaps: 19

US-10-086-464-2 (1-647) x US-10-425-114-7339 (1-2655)

Qy 2 SerSerAlaProSerProGlyThrGlySerProSerProSerProSerAsnSer----- 19  
Db 98 TCATCTCTCCACCA---AACACTCCACCCACATCACTCCAGATCTCTCAGCCT 154  
Qy 20 -----ThrThrThrThrProProProAlaSerAlaPro---ProProThrThr 34  
Db 155 AATCAACACAAACACAAACACAAACCAATCCCTGCTCTCATCTCTCTCTGAC 214  
Qy 35 ProSerSerProPro-----ProProSerThrThrThrThrThrThrThrThrThrThr 52  
Db 215 CTTTCTGCTCCACCT 274

Qy 53 ArgSerThrPro-----SerAlaProProProSerProProThrProSerThrPro 69  
Db 275 TCACCTGCTCTCTCCACCATCATCTACACCAACCACTTTGCCCTCCATCACCACCT 334  
Qy 70 GlySerProProProLeuPro-----GlnPro 78  
Db 335 GATTCTCCACCACTTGGCCACTGTCATCTCAACCCACCACAGTTACAACAATCTCCCCC 394  
Qy 79 SerProProAlaProThrThrThrGlySerProProAlaProValThrProProThrArg 98  
Db 395 TCACCAACAGTTACAACATCTCCCTCTCACACAGTACCACATCCCTCTCTCCAGCT 454  
Qy 99 AsnProProProSerValPro-----GlyProProSerAsnProSerArgGluGly 116  
Db 455 GAAACTCCACCTCTCCCTTCCAAATTTGCTCTCCACCACTCTCCACCCAGC-----GGT 505  
Qy 117 SerProArgPro-----ProSerSerPro 124  
Db 506 TCCCTCTCTCTCAATCACCTCCCAACAATTCACCTCTCTCTCGACCGATTTCCGCA 565  
Qy 125 SerProProSerProSer----- 130  
Db 566 TCTCTCTCTCTCGGCCAAATTTTCCAAAGGCCCAACCACTAGAACTCTCTCCAGAGAA 625  
Qy 130 ----- 130  
Db 626 GAAATCCACCAAAAACCTACTCTCTTACATGCATCTCTCCATCAGTTTCTGAAACTCCT 685  
Qy 130 ----- 130  
Db 686 CCTAAACCTCTCTCTGATGTTCCCTCTCCATCCACATTCCTCTCACTCTCTCTTCA 745  
Qy 130 ----- 130  
Db 746 GACCTTTCAGATCTTCACTCTCAGCTTCTTTGCTGATCCCTCCCACTAATAAACAAGTG 805  
Qy 130 ----- 130  
Db 806 GTGGGGGTCCGAAGGTGTCGTACCTCTCTTCCAACTGAGAAACCACTAGTACCT 865  
Qy 131 -----SerAspGlyLeu 134  
Db 866 ACTAATGATGGTACTTAACAGTATGTCTTCAAAACACACACCTTCCATCTCTGAGGGTTG 925  
Qy 135 SerThrGly-----ValValValGlyIleAlaIleGlyGlyValAlaLeuValIle 152  
Db 926 AGCACTGGAGATCTGTGGCTATTGGAAATGTGTTGTTTATTGTTCTCAGCCTCTCT 985  
Qy 153 ValThrLeuIleCysLeuLeuCysLysLysLysLysLysLysLysLysLysLysLys 172  
Db 986 GTTATGGCTGTGTTTGGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1045  
Qy 173 TyrValProProProProProProGlyProLysAlaGlyGlyProTyrGlyGlnGln 192  
Db 1046 TATGCTGCTCTCTCTCCA----- 1063  
Qy 193 GlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuPro 212  
Db 1064 -----TTTACCTCATCCCAATTCAGGTACCTTATTC 1096  
Qy 213 ProProProLysAlaProSerProProArgGlnProProProProProProPhe 232  
Db 1097 TTGAGGCGGAGTCTCGGCC-----AACTTT 1123  
Qy 233 MetSerSerSerGlySerAspTyrSerAspArgProValLeuProProProSerPro 252  
Db 1124 TTAGGTAGTGGCTGTGGTAGTGATTTT-----GTATATTCTCCATCAGAGCCT 1171  
Qy 253 GlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaAArgAlaThr 272  
Db 1172 GGTGGTGA---AGTAGTTCAAGATCATGTTTACATATGAAGAACTTATTTCAAGCTACA 1228  
Qy 273 AsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGly 292

Db 1229 AATGGGTTTCAGCAAAATTTGTTGGAGAAGGTGGATTGGCTGTGTTTATAAAGT 1288  
QY 293 ValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGly 312  
Db 1289 TTGCTGATAGATGGAAGAGAGTAGCTGTGAAACAGCTCAAAATTTGGTGGTGGCAAGG 1348  
QY 313 GluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisHisArgHisLeuVal 332  
Db 1349 GAACCGCAATTCAGGCGCAGAGTTCAGATTATTTAGCCGTGTACATCATCGTTCGTGTT 1408  
QY 333 SerLeuValGlyTyrCysTleAlaGlyAlaLysArgLeuLeuValTyrGluPheValPro 352  
Db 1409 TCTTTAGTTGGTTACTGTATATCCGAGATTCAGATTGCTTTATATAGACTATGTTCC 1468  
QY 353 AsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSerThr 372  
Db 1469 AAGGATCTCTTCATTACCATCTCCACGGTGAATAATACACCAAGTTCTAGATTGGCCTACC 1528  
QY 373 ArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsn 392  
Db 1529 AGAGTCAAGTGTCTGCTGGTGCAGCTCGTGAATAGCTTACTTGCATGAAGACTGTCTAT 1588  
QY 393 ProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412  
Db 1589 CCACGCATTTATTCATCGAGATATTAAAGTCATCAAAATCCTTACTTGTATCTCAACATATGAA 1648  
QY 413 AlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer 432  
Db 1649 GCTCAAGTTTCGGACTTTGGGCTTCGAAAATTTGGCATTTAGATTCAAAATACACATGTA 1708  
QY 433 ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 452  
Db 1709 ACACGTGTAAATGGGAACCTTTGGGTACATGCGCAGCAATATGCCAAGTGGAAAACTT 1768  
QY 453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 472  
Db 1769 ACTGAAAGTCTGATGATATTTCTTTGGGGTGTGCTTTGGAGCTAATTACAGTTCGG 1828  
QY 473 ArgProValAspAlaAsnValTyrValAspSerLeuValAspTrpAlaArgPro 492  
Db 1829 AAGCTGTAGATGATCTCAACCAATTTGGTATGAGAGCTGGTTGAATGGCTGCACCT 1888  
QY 493 LeuLeuAsnArgAlaSerGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsn 512  
Db 1889 CTGTTGACGGAAGCACTTGACAAATGAGGACTTTGAAATTTTGGTGGATCCAAGACTGGGG 1948  
QY 513 AsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532  
Db 1949 AAGAACTACGATAGAAATGAATGTTTCGATGATCGAGGCTGCCGCGCTGTGTACGC 2008  
QY 533 HisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552  
Db 2009 CACTCATCGGTGAAGAGACCAACGATGAGTCAAGTGGTGGAGCTTTAGATTCTTGAT 2068  
QY 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr 572  
Db 2069 GAGTTTACGGATCTCAATAACCGAATGAACCGGACACAGTTCGGTG----- 2116  
QY 573 GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys 592  
Db 2117 -----TTGATTCCGGCAGCAATCTGCACAAATCAGAAATGTTTAGGAGG 2161  
QY 593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp 612  
Db 2162 ATGGCTTTTGGAGGCCAAGAT-----AGTTCGGTTCCTTCAATGAGTCTCAGAGTAGC 2215  
QY 613 Tyr-GlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluMetGluWe 632  
Db 2216 TGGAGGAGTA-----GAGATCATCAGCAACCACTGCTGCTTC 2251  
QY 632 tgllylsilleysArgThrGlyGlnGly 641

Db 2252 TCCCAAAATAAAA---CTGGGCATGGGA 2276  
RESULT 14  
US-10-425-114-14127  
; Sequence 14127, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 14127  
; LENGTH: 2190  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB148-034-D5\_FLI  
US-10-425-114-14127  
Alignment Scores:  
Pred. No.: 1.09e-84 Length: 2190  
Score: 1437.50 Matches: 331  
Percent Similarity: 61.90% Conservative: 80  
Best Local Similarity: 49.85% Mismatches: 167  
Query Match: 41.63% Indels: 87  
DB: 13 Gaps: 19  
US-10-086-464-2 (1-647) x US-10-425-114-14127 (1-2190)  
QY 1 MetSerSerAlaProSerProGlyThrGlySerProPro----- 13  
Db 2 ATGATGCTTCTCTCCCTCCCTCATTCCTGCTCGCCCTCCATTCACCTTCGACGCCCTT 61  
QY 14 SerProProSerAsnSerThrThrThrProProProAlaSerAlaProPro----- 31  
Db 62 TCTCTTCTCCTCTCTCAATCCTGTCAGCGAAAGGAAAGGGTGTGACACCGTA 121  
QY 31 ----- 31  
Db 122 AGAGCTATTTAGCAATAATACATACATAGCTATATACCTCTTTTAGGTCAAGAGGTGGAGT 181  
QY 32 ---ProThrThrProSerSerProProProSerThrThrPro----- 45  
Db 182 TCTCTTATTTCGGAGGAGAGCTCCATCTCTCATTTACAGACCAACCAATAAATGGCC 241  
QY 46 ThrSerProProProSerSerArgSerThrProSerAlaProProProProProThr 65  
Db 242 TCTAGCCCAAGTCTCTGCCATCATCTCCCAAGAGTCCGCCCCACCAACAGAGCATCA 301  
QY 66 ProSerThrProGlySerProProProProGlnProSerProProAlaProThrThr 85  
Db 302 GCTTTGGGCGCAAAATCATCTTCA-----AAGTCACCTCGCCACCAAGGGTGAATCA 355  
QY 86 ProGlySerProProAlaProValThrProProThrArgAsnProProProSerValPro 105  
Db 356 ---AATTTCATCTCCAGTCCCTCCATCGAAATCGCTCTCCATCCGCCCGCCACCAAG 412  
QY 106 GlyProProSerAsnProSerArgGlyGlySerProArgProProProSerProSer 125  
Db 413 AAAAGTGTAGTATCTAGTCTTAAGATGGGAGCAGCAAGAGTCTCTTCATCGTCT 472  
QY 126 ProProSerProSerAspGlyLeuSerThrGlyValValValGlyIleAlaIleGly 145  
Db 473 TCATCGTCGGAATAACACCGCTGTGATCACCAGCGTGTGCTTGGGGTG---GTGGGC 529

Qy 146 GlyValAlaLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArg 165  
Db 530 TTGCTCTGCTCTTGTCCATCGTGGCGTGTGTGTCGCGCAAGAAGAAAGAGT 589  
Qy 166 ArgAspGluGluAspAlaTyrValProProProProProProGlyPro 182  
Db 590 -----CCGCCCAATGAACATGCCCTTCTACACC 619  
Qy 183 ---LysAlaGlyGly-----ProTyrGlyGlyGlyGlyGlyGlyGln 198  
Db 620 GACGAGAAGCAATGTATTACCCCAACGCTGTGCGCCCTATGTGCACCAATAT 679  
Qy 199 AsnAlaThrProProSerAspHisValValThrSerLeuProProProProLeuAlaPro 218  
Db 680 GGCAGCAAC-----GGCAGCATCCCTCCGCCGGG-ATGGCACCA 717  
Qy 219 SerProProArgGlnProPro-ProProProProProPheMetSerSerGlyGln 238  
Db 718 TCACGGCGCGGGAACCGCTGTGCCCATCGCTCGGCTCGATGGCGGCCCTTGAGCGG 777  
Qy 238 ySerAspTyrSerAsp-----ArgProValLeuProProProSerProGlyLeuVa 255  
Db 778 CGAGATGACTCGTCCGCCGCCCGCCCGGCTTCCCTTCCCGCCACCTTCCCGCAAGCTGGC 837  
Qy 255 lLeuGlyPheSerLysSerThrPheThrTyrGluLeuAlaAlaArgAlaThrAsnGlyPh 275  
Db 838 GCTGGGTTCCTCAAGAGCTCTTCTCGTACGAGGAGCTAGCGCGAGCGTCCGGCTT 897  
Qy 275 eSerGluAlaAsnLeuLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyValLeuPr 295  
Db 898 CTCGTCGCGCAACCTGCTGGCGGCGAGCGGCTTCCGGGTACGTGTACAGGCGTGTCTCGC 957  
Qy 295 o---SerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlyGlyGluAr 314  
Db 958 CGGCGAGCGGAGAGGTGGCGGTGAACGACTCAAGTCCGGAGCGCGGAGGGAGCG 1017  
Qy 314 gGluPheGlnAlaGluValGluIleSerArgValHisHisHisHisLeuValSerLe 334  
Db 1018 CGAGTTCAGCGGAGGTGGAGATCATCAGCGCGGTGCACCCACCTGCACCTGGTGTCTGCT 1077  
Qy 334 uValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAs 354  
Db 1078 CGTGGCTACTGTCATCGCGGCAACACGCGCATGCTGTCTACGAGTTCGTGGCCAAAC 1137  
Qy 354 nAsnLeuGluLeuHisLeuHisGlyGlyGlyArgProThrMetGluThrSerThrArgLe 374  
Db 1138 CACGCTGGACCATCTGTAGCCAGGAGCGGCCCTGTCTATGAGCTGGACACCCGCT 1197  
Qy 374 uLysIleAlaLeuGlySerAlaLysGlyLysLeuSerTyrLeuHisGluAspCysAsnProLy 394  
Db 1198 GAAGATCGCGCTCGGCTCCGCCAAGGGCTCGCTTACCTCCACGAGGACTGCCATCTCG 1257  
Qy 394 eIleIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLy 414  
Db 1258 GATCATCCCGCGACATCAAGGCGCGCAACATCTGTGTGACCAACACTTCGAGGCCAT 1317  
Qy 414 eValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrAr 434  
Db 1318 GGTGGCGACTTCGGGCTGGCAAGCTCACCGCGAGCAACCAACGACGCTCTCCAGCG 1377  
Qy 434 gValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGl 454  
Db 1378 CGTCATGGCACCTTCGCTACCTTGGCGCGGAGTACGGCTCCAGCGGCAAGCTCACGGA 1437  
Qy 454 uLysSerAspValPheSerPheGlyValValLeuLeuLeuLeuThrGlyArgArgPr 474  
Db 1438 CAGTTCGGACGCTCTCTTCGCGCTCATGCTGTGGAGTACTCATCGGGCGCGCCGCC 1497  
Qy 474 oValAspAlaAsnValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLe 494  
Db 1498 CATGATACCAACCAAC---TACATGGAGGAGCAGCTCGTGGACGTGGCGCGCGCTGCT 1554  
Qy 494 uAsnArg-----AlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAs 512

Db 1555 GAGCGCGCGTGGCGGGGAGACGGC---TTGCGCGAGCTCGTCCACCCCTCTGGG 1611  
Qy 512 nAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAlaAlaCysValar 532  
Db 1612 AGGCGAGTACTCGCTGTCGAGGTGGAGCGCTTGGCGGCTTGGCGCGCGCACACCGG 1671  
Qy 532 gHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVa 552  
Db 1672 CCACTCGGCCAAGCGCGCCCAAGATGAGCCAGATCGTGGCGCGCTGGAGGGGACGC 1731  
Qy 552 lSerLeuSerAspLeuAsn---GluGlyMetArgProGlyGlnSerAsnValTyrSerSe 571  
Db 1732 ATCGCTGGAGGATCTGCACGAGGACGGCGGCAAGCCCGGCGAGCGTGTCTTCTTC- 1789  
Qy 571 rTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheAr 591  
Db 1790 ---GGCGGTGGCTCCGAC-----AACATCTCCCGCTTAAG 1821  
Qy 591 gLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSe 611  
Db 1822 GCAGCTCGCTTCGACAGCGCGGACTACTCG-----GACTACAGCAGCGACTCCTC 1872  
Qy 611 rAspTyrGly 614  
Db 1873 CACCGTGGT 1882

RESULT 15  
US-10-425-114-6300  
; Sequence 6300, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 6300  
; LENGTH: 1500  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700560157\_FLI  
US-10-425-114-6300

Alignment Scores:  
Pred. No.: 8.16e-85 Length: 1500  
Score: 1437.00 Matches: 284  
Percent Similarity: 74.27% Conservative: 45  
Best Local Similarity: 64.11% Mismatches: 80  
Query Match: 41.62% Indels: 34  
DB: 13 Gaps: 7

US-10-086-464-2 (1-647) x US-10-425-114-6300 (1-1500)

Qy 185 GlyGlyProTyrGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSer 204  
Db 11 GGAGGTCCACATGGG----- 25  
Qy 205 AspHisValThrSerLeuProProProProProProProProProProProProProPro 224  
Db 26 GACCATGTTGTG---AGGATGCACAAATGGATGGTCTCTGGTGGCGGAGGTGGGT 82  
Qy 225 ProProProProProProPheMetSerSerGly-----GlySerAspTyrSer 242  
Db 83 GCACCAACCACTCTCTCCATGATGATGAGTAGTGTCTGAGTTTAGTCTCCAATTACTCG 142

Qy 243 AspArgProVal----LeuProProSerProGlyLeuValLeuGlyPheSerLysSer 261  
 Db 143 ACGGACCGCCCGCTTGGCTCTCCCTCACCACAACTTGTCTGGGCTCAAAAGGGGC 202  
 Qy 262 ThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 281  
 Db 203 ACCTTCACTTACGAGGAATTAGCAGCTGCCACCAACGGATTCAATGATGCAAAATTGATA 262  
 Qy 282 GlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAla 301  
 Db 263 GGACAGGTGATTCGGCTATGTCCATAAGGGTGTGTTCCTAGTGGAAAGAAAGTGGCA 322  
 Qy 302 ValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGlu 321  
 Db 323 GTTAAGAGCTTAAAGCAGGTAGTGGCCACAGAGAGCGAGAGTTCCAAAGCTGAGATTGAC 382  
 Qy 322 IleIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGly 341  
 Db 383 ATCATTAGCCGGCTCATCGCCATCTGTGTCACTTGTGGATCAATTCTGGT 442  
 Qy 342 AlalysArgLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHis 361  
 Db 443 GGCCAGAGGATGTTGGTCTATGAATTTATTCCTCAATAACACATTGGAAATATCACCTCCAT 502  
 Qy 362 GlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAla 381  
 Db 503 GGAAGGGTAGGCCTTACCATGCTGGCCAACTAGAATCGGATTCGAATAGGATCCGCT 562  
 Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLys 401  
 Db 622 AAAGGGCTTGCTTATCTCATGAGGACTGTCTCTCGTATTATCCATCGCGATATCAAA 421  
 Qy 402 AlaSerAsnIleLeuIleAspPheLysPheGluAlalysValAlaAspPheGlyLeuAla 421  
 Db 682 GCTGCAATGCTTATTTATGATGATAGCTTCGAGCAAGGTTGCTGATTTGGATTGGCT 682  
 Qy 422 LysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyr 441  
 Db 742 AAGTTGACTACTGATAATAATACTCATGTATCGACTCGTGTCTGGGAACTTTCCGGGTAC 742  
 Qy 442 LeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPhe 461  
 Db 802 CTAGCCCTCGAATATGCAATCAAGTGGAAATTGACAGAGAAGTCTGACGTTTTCTCATTT 802  
 Qy 462 GlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyr 481  
 Db 859 GGAGTCATGCTATTGGAACTCATAACTGGGAAGCGACCTGTTGATCACACAAATGCC--- 859  
 Qy 482 ValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAla---SerGluGln 500  
 Db 919 ATGGACGACAGCCTAGTAGACTGGGCTCGACCCTTCTGACTCGTGGACTAGAGGAGGAT 919  
 Qy 501 GlyAspPheGluGlyLeuAlaAspAlalysMetAsnAsnGlyTyrAspArgGluGluMet 520  
 Db 979 GGCAACTTTGAGAGTCTGTTGATGCATTTTGGAGGAAACTAGATCCTCAAGAACTG 979  
 Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArg 540  
 Db 1039 TCAAGATGGCAGCTTGTGCTGCCGTAGCATTCTGTCATTCTGCCAAAAACGTCCTCAAAA 1039  
 Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560  
 Db 1040 ATGAGCCAGATTGTAAGAATATTGGAGGAGATGTCTCCTGATGACTTGAAGAATGGG 1099  
 Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlySerThrAspTyrAspSer 580  
 Db 1159 ATTAGCCGGGCGCAAAATGTTGCTTACAACTCTTCATCTAGCTCAGATCAGTATGACACA 1159  
 Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600  
 Db 1219 ATGCAATACAATGCTGATATGACAGAAGTTCAGAAAGCGCAGTGTCTTCTAATAGCGAGGA 1219

Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620  
 Db 1220 TTGGCACCAGC-----AGTGGCTCC 1240  
 Qy 621 SerSerGlu 623  
 Db 1241 AGTGGTGAA 1249

Search completed: April 25, 2004, 02:33:32  
 Job time : 696 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2004, 20:49:18 ; Search time 5877 Seconds  
(without alignments)  
4771.643 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSGTGTGPPSPNST.....REMEMGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=x1p  
-Q=/cgn2/1/USPTO\_spool\_p/US10086464/runat\_11052004\_121814\_7841/app\_query.fasta\_1.839  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10086464 @CGN\_1\_1\_5265 @runat\_11052004\_121814\_7841 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_em.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pi.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pin.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3453	100.0	1944	6	AX088876 Sequence
2	3453	100.0	1944	6	AX825703 Sequence
3	3453	100.0	2189	6	AX825705 Sequence
4	3453	100.0	2189	6	AY028699 Brassica
5	2936.5	85.0	1959	6	AX825738 Sequence
6	2936.5	85.0	1959	6	BT008400 Arabidops
7	2936.5	85.0	2098	8	AY128792 Arabidops
8	2936.5	85.0	2116	8	AY056788 Arabidops
9	2936.5	85.0	2188	8	AY059901 Arabidops
10	2936.5	85.0	2190	8	AY093065 Arabidops
11	2934.5	85.0	1959	8	BT008409 Arabidops
12	2934.5	85.0	2257	8	AF370509 Arabidops
13	2885	83.6	2324	8	AY089024 Arabidops
14	2622	75.9	79706	8	AB020746 Arabidops
15	1841.5	52.3	2307	8	AK103247 Oryza sat
16	1821.5	52.8	1812	6	AX825735 Sequence
17	1812	52.5	80393	8	AP000382 Arabidops
18	1748.5	50.6	1902	6	AX088882 Sequence
19	1748.5	50.6	1902	6	AX825712 Sequence
20	1748.5	50.6	1939	6	AX088881 Sequence
21	1748.5	50.6	1939	6	AX825711 Sequence
22	1746	50.6	2013	8	BT005955 Arabidops
23	1746	50.6	2320	8	AK118488 Arabidops
24	1687	48.9	2025	6	AX088885 Sequence
25	1687	48.9	2025	6	AX825715 Sequence
26	1687	48.9	2104	6	AX088884 Sequence
27	1687	48.9	2104	6	AX825714 Sequence
28	1668	48.3	1515	6	AX825736 Sequence
29	1637.5	47.4	1488	6	AX825737 Sequence
30	1574.5	45.6	1185	8	AY224526 Oryza sat
31	1567.5	45.4	2364	8	AK064397 Oryza sat
32	1567.5	45.4	2371	8	AK102832 Oryza sat
33	1507	43.6	82356	8	AC006135 Arabidops
34	1507	43.6	92861	8	AC006439 Arabidops
35	1480.5	42.9	125021	8	AC007504 Arabidops
36	1479.5	42.8	114367	8	ATF10M10 Arabidops
37	1479.5	42.8	125502	8	ATT4L20 Arabidops
38	1479.5	42.8	195921	8	ATCHRIV81 Arabidops
39	1459.5	42.3	2158	8	AY113877 Arabidops
40	1459.5	42.3	3064	8	AY035076 Arabidops
41	1456	42.2	2046	8	AY113039 Arabidops
42	1456	42.2	2616	8	AF424623 Arabidops
43	1456	42.2	2735	8	AY075681 Arabidops
44	1434.5	41.5	120465	2	AC137001 Oryza sat
45	1431	41.4	3070	8	AK110196 Oryza sat

ALIGNMENTS

RESULT 1

AX088876  
 LOCUS AX088876 1944 bp DNA linear PAT 17-MAR-2001  
 DEFINITION Sequence 1 from Patent WO0114563.  
 ACCESSION AX088876  
 VERSION AX088876.1 GI:13397639  
 KEYWORDS

SOURCE Brassica napus (rape)

ORGANISM Brassica napus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1

AUTHORS Goring, D. and Silva, N.

TITLE Proline-rich extensin-like receptor kinases

JOURNAL Patent: WO 0114563-A 1 01-MAR-2001;

Goring, Daphne (CA) ; Silva, Nancy (CA)

FEATURES Location/Qualifiers

source

1..1944  
 /organism="Brassica napus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3708"

CDS

1..1944  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC34703.1"  
 /db\_xref="GI:13397640"  
 /translation="MSSAPSGTGTGSPSPSPSNSTTTTPPASAPPPTTPSPSPSPPTT  
 PTPSPSPSTPSAPPSTPSPPSPPLPQSPAPPTPGSPSPPTGSPSPAPVPTPTNPP  
 SVGPSPNPSREGSPSPSPSPSPSSDGLSTGVVGTAGTGVALLVITLLCLC  
 KKRRDEEDAIYVPPFPKAGPIGGQQQROQNPATPSDHWVLSLPPPKAP  
 SPQPFPFPFPSPSPSSGSDSDRPLPPSPGLVGFSGSKFTFYEEELARATNPS  
 EANLGGGGYVHKGLVPSGKEVAVKLVKVGQGEREFQAEVEIIISRVHRLHVL  
 VGYICAGKRLVVEFVNNLELHLEGRTMWSRLKIALGSAKGLSYLHEDCN  
 PKIHRDIKASNLILDFKFAKVADFGLAKIADTNTHVTRVMGTFGYLAPEYVAGS  
 KLTEKSVFSGVLLLELIGRRPVDANNVYVDDSLVDWAPLILNRASEQDFGLAD  
 AOMNGXDREEMAKWACAAACVHRHARRRMSQIVALEGNVSLDLNEGMRPGQS  
 NVYSYSGSDTDYSSDQMDKKFRKMALGTQYENATGEVSNPTSDYLYSGSSSEG  
 QTTREMEMGKIKRTGGYSGPSL"

ORIGIN

Alignment Scores:  
 Pred. No.: 4,02e-76 Length: 1944  
 Score: 3453.00 Matches: 647  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-086-464-2 (1-647) x AX088876 (1-1944)

Qy 1 MetSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20  
 Db 1 ATGTCCTCGGCGCGTCTCGGGGACTGGTTCCGCTCCATCTCCACCATCAAACTCCACA 60  
 Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProProSerProProPro 40  
 Db 61 ACCACCACTCTCTCCAGCTTCGGCTCTCTCTCCACCACTCTCTCTCTCTCTCTCGCG 120  
 Qy 41 ProSerThrIleProThrSerProProSerProSerArgSerThrProSerAlaProPro 60  
 Db 121 CCATCCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
 Qy 61 ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80  
 Db 181 CCATCTCCACCACTCCATCTACGGCGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCT 240  
 Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100  
 Db 241 CCCGCTCCAACTACGCGCGGATCTCTCCACCCGACCTGTGTACTCTCTCTCTCTCTCTCT 300  
 Qy 101 ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro 120

Db 301 CCACCTTCAGTCCAGGACCAACCGTCCAACTCTCACGCGAAGAGGATCTCCTCGACCT 360  
 Qy 121 ProSerSerProProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140  
 Db 361 CCATCTTCTCCTCGCGCGCT 420  
 Qy 141 GlyTleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160  
 Db 421 GGAATCCCACTCGGAGGAGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
 Qy 161 LysAlaLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProProPro 180  
 Db 481 AAGAAGAAACGACGAG 540  
 Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnIntrpA-gInGlnAsnAla 200  
 Db 541 GGTCCCAAGCGGAGGACCTTACGGTGGACAGCAACAATGGCGGCAACAAACGCA 600  
 Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProLysAlaProSerPro 220  
 Db 601 ACACCACTCAGATCATGTGCTGACGTCACTACCAACCAACCAACCAACCAACCAACCA 660  
 Qy 221 ProArgGlnProProProProProProProPheMetSerSerSerGlyGlySerAsp 240  
 Db 661 CCACGGCAACTCTCTCCACTCCACCAACCGCTTTCATGACACAGCGGCGGCTCCGAC 720  
 Qy 241 TyrSerAspArgProValLeuProProProProProGlyLeuValLeuGlyPheSerLys 260  
 Db 721 TACTCGGACCGTCCAGTTCTTCTCTCCACCGTCTCCAGGCTTGTGTAGGCTTCTCCAAA 780  
 Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280  
 Db 781 AGCACTTTTCATACAGAGAGAGTACGTAGAGCCACCAATGGTGTCTCCGAGCGCAACTG 840  
 Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisGlyValLeuProSerGlyLysGluVal 300  
 Db 841 TTAGGACAGCGCGTTCGGTTAGTCTGCACNAAGTGTGTTCCTAGTGTGGAGAAAGATT 900  
 Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320  
 Db 901 GCTGTGAAGCAGTTGAAAGTTGGAGTGGTCAGGAGAGAGAGAGTTTCAGGACAGAGTT 960  
 Qy 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340  
 Db 961 GAGATCATACAGAGAGTTCCACCAAGAGTCTGGTGTCTCTTGTGTGTATGTGATCGCC 1020  
 Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360  
 Db 1021 GGTGCCAAAGATTGCTTGTCTATGAGTTTGTCTTAAACAACATCTCGAGCTTCACCTC 1080  
 Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380  
 Db 1081 CATGCGAGGAGCGGCTTCAATGGAATGGAGCACCAAGATTGAAGATTGTCTTTGGATCT 1140  
 Qy 381 AlalaGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400  
 Db 1141 GCTAAAGAGCTTCTTATCTTCAAGAGATTGCAATCTCTAAATCATTTACCGGTATATC 1200  
 Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420  
 Db 1201 AAGCTTCAACATATTGATAGATTTCAGTTTGAAGTAAAGTTGTGATTTTGGTCTT 1260  
 Qy 421 AlalaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440  
 Db 1261 GCTAAGATTGCTCTGATACAAACACGATGTATCAACACGCTGTGATGGGAACCTTTGGG 1320  
 Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460  
 Db 1321 TACTTGGCTCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTCA 1380  
 Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480  
 Db 1381 TTTGGCGTTGTCTTTTGGAGCTCATTTACTGGACGTGACCGCTTGATGCCCAACAATGTC 1440

Qy 481 TyrValAspSerLeuValAspTrrpAlaArgProLeuLeuAasnArgAlaSerGluGln 500  
 Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGCGACGACCATTTGCTTAACCGAGCATCTGACGAA 1500

Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAenAenGlyTyrAspArgGluGluMet 520  
 Db 1501 GGAGACTTTGAGGGTTAGCTGATGCCAAGATGAATAATGGGTATGACAGAGAGGATG 1560

Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaAArgAtqProArg 540  
 Db 1561 GCTCGCATGGTGTGCTGGCGCTGTGTTGCCATTGAGCTGCGCGAGACCTCGC 1620

Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAasnGluGly 560  
 Db 1621 ATGAGCCAGATTGTCGGTGGCTAGNAGGAATGATACACTGTCCAGATCTTAACGAAGG 1680

Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580  
 Db 1681 ATGAGACGAGTCAAGCAATGTATACAGCTCATACGAGGAGGAGACCGATTATGACTCG 1740

Qy 581 SerGlnTyrAenGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600  
 Db 1741 AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800

Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620  
 Db 1801 AACGCCAGGTGAGTACAGTAAATCCGACCACTGACTATGGACTGTACCCGCTGTGGTCA 1860

Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640  
 Db 1861 AGCAGCGAGGCCAAACCAACGAAATGAGATGGGAAGATTAAGAGAACGGGTGAG 1920

Qy 641 GlyTyrSerGlyProSerLeu 647  
 Db 1921 GGTTATAGTGACCTTCTCTT 1941

RESULT 2  
 AX825703  
 LOCUS AX825703 1944 bp DNA linear PAT 11-DEC-2003  
 DEFINITION Sequence 1 from Patent WO03072763.  
 ACCESSION AX825703  
 VERSION AX825703.1 GI:39751230  
 KEYWORDS  
 SOURCE Brassica napus (rape)  
 ORGANISM Brassica napus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1  
 Goring, D., Silva, N. and Haffani, Y. Z.  
 Increasing plant seed production  
 Patent: WO 03072763-A 1 04-SEP-2003;  
 JOURNAL Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)

FEATURES  
 source  
 1. .1944  
 /organism="Brassica napus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3708"  
 1. .1944  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAE84846.1"  
 /db\_xref="GI:39751231"  
 /translation="MSSAPSGTGSPPSPSNSTTTTPPPASAPPTTPSPSPSPSTI  
 PTPSPSPSTPSAPPPTPTPTGPPPLPQSPAPPTPGSPAPPTPGSPAPVTPPTRNPP  
 KVGPPNSPREGSPSPSPSPSSDGLSTGVVVGIAIGGVALLVITVILCLIC  
 SKRRRREDAIYVPPPPPGPKAGGPGVQOQOQOONATPPSDHVVVTLVPPPKAP  
 SPQPQPPPPPPPMSSSGGSDYSDRVLPPSPGLVLPFSKSTFTVEELARATNGS  
 EANLGGGFGYVHKVLPSGKEVAHVQLKVGSCQGEREQAEVIELIRVHRHLVSL  
 VGYCAGAKRLLVTFVFNNNLEHLHDEGRPTNENWTRLKIALGASGLSYLHEDCN  
 PKIHRDKASNLIDFKFAKADVFLAKIADSTNTHVSTVMTGTFGLAPYAAAG  
 KLTEKSDVFGVGVLLLELTGRRPVDANNVVDLSLVDWARPLNLRASEQGDPEGLAD

CDS

## ORIGIN

Alignment Scores:  
 Pred. No.: 4.02e-76 Length: 1944  
 Score: 3453.00 Matches: 647  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-086-464-2 (1-647) x AX825703 (1-1944)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProSerAsnSerThr 20  
 Db 1 ATGTCTCTCGGGCGGCTCTCGGGGACCTGGTTCGGCTCCATCTCCACCATCAAACTCCACA 60

Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40  
 Db 61 ACCACCACTCTCTCTCCAGCTTCGGCTCTCTCTCCACACACACCTTCTCTCTCCGCG 120

Qy 41 ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaProPro 60  
 Db 121 CCATCCACTATTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

Qy 61 ProSerProProThrProSerThrProGlySerProProProProProGlnProSerPro 80  
 Db 181 CCATCTCCACCACTCCATCTAGCGGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA 240

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100  
 Db 241 CCGCTCCAATACTAGCGCGGATCTCCACCGCGACCTGTACTCTCTCTCTCTCTCTCTCT 300

Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120  
 Db 301 CCACCTTCAGTCCCGAGGACCGGTCCATCTCTTACGCGAAGGAGGATCTCTCTCGACCT 360

Qy 121 ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140  
 Db 361 CCATCTCTCTCTCGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGT 420

Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160  
 Db 421 GGAATGCCATCGGAGGAGTGGTCTGCTGTGTGATAGTACTGATCTGATTGTTCTCTCTGT 480

Qy 161 LysLysLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProProPro 180  
 Db 481 AAGAAGAAACGACGAGAGACGAAAGATGCTTACTATGTCTCTCGCCACCTCTCTCTCTCT 540

Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnTyrArgGlnGlnAsnAla 200  
 Db 541 GGTCCCCAAACCGGAGGACCTTACGGTGGAGACAGCAACAATGGCGGCAACAAACGCA 600

Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProLysAlaProSerPro 220  
 Db 601 ACACCCCGTCAGATCATGTCTGTCGTCATCTACCACCACTAAGGCTTCCATCTCCA 660

Qy 221 ProArgGlnProProProProProProProPheMetSerSerSerGlyGlySerAsp 240  
 Db 661 CCACGGCAACTCTCTCCACCTCCACCGCTTTCATGAGCAGCAGCGCGGCTCCGAC 720

Qy 241 TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys 260  
 Db 721 TACTCGACCGCTCAGTTCCTCTCCACCGCTCTCCAGGGCTTGTGTAGGCTTCTCCAAA 780

Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280  
 Db 781 AGCACTTTCATACAGGAGGAGCTAGCTAGACCACTAATGTTCTCTCGAGGCGAATCTG 840

Qy 281 LeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300

AKNNNGYDREEMARMVACAAACVRHSARRRRPMRSQIVRALEGNVLSLSDLNENGMRRPGOS  
 NVTSSYSGSDYDSSQVNEDEKFRKMACTOBYNATGEYSNPTSDYGLYPSGSSESG  
 QTTREMEMGKI KRTGCGYSGPSL"

Db 841 TTAGCAAGGGGGTTTCGGTTACGTGCACAAAGGTGTGTTGCCCTAGTGGGAAGAGTT 900  
Qy 301 AlaVallysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320  
Db 901 GCTGTGAAGCAGTTGAAAGTTGGGAGTGGTTCAGGGAGAGAGGAGTTTCAGCGCAGAGGTT 960  
Qy 321 GluIleIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340  
Db 961 GAGATCATCAGCAGAGTTTACCACAGGATCTGGTGTCTCTTGTGGTATTTGATCGCC 1020  
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360  
Db 1021 GGTGCCAAAGATTGCTTGTCTATGAGTTTGTCTTAAACAATCTCGAGCTTCACCTC 1080  
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380  
Db 1081 CATGGCAGGAGCGGCTTACAATGGAATGGAGCACAGATTGGAAGATTGCTCTTGGATCT 1140  
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400  
Db 1141 GCTAAAGACATTTCTTATCTTCATGAAGATTGCAATCTTAAATCATTCACCGTGATATC 1200  
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420  
Db 1201 AAGGCTTCAACATATGATAGATTCAAGTTTGAAGCTAAGTTGCTGATTTGGTCTT 1260  
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440  
Db 1261 GCTAAGATTGCTTCTGTATACAAACAGCATGTATCAACACGTGTGATGGGAACCTTTGGG 1320  
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460  
Db 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGCTTTCTCA 1380  
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480  
Db 1381 TTTGGCGTTGTCTTTGGAGCTCATTTACTTGAGCTGCAGCTGACCGTTGATGCCAACAATGTC 1440  
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500  
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTGCTTAAACCGAGCATCTGAGCA 1500  
Qy 501 GlyAspPheGluGlyLeuAlaAspAlalysMetAsnAsnGlyTyrAspArgGluMet 520  
Db 1501 GGAGACTTTGAGGGTTTGTAGCTGATGCAAAAGATGAATAATGGGTATGACAGAGAGAGATG 1560  
Qy 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540  
Db 1561 GCTGCAATGGTTGTGTGCTGGCGTTGTGTTCGCCATTACGCTCGCGCAGACCTGCG 1620  
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560  
Db 1621 ATGACCCAGATTGTGCGTGGTTAGAGGAATATGATCACTGTACAGATCTTAAACGAAGG 1680  
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580  
Db 1681 ATGACCCAGGTCAAAGCAATGTATACAGCTCATACGGAGGAAGCAGGATATGACTCG 1740  
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600  
Db 1741 AGCCAGTACATGAGACATGAGAGAGTTTAGAANAATGGCACTTGGAACTCAAGAGTAC 1800  
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620  
Db 1801 AACGCCACGGGTAGTACAGTAATCCGACCAGTACTATGGACTGTACCCCTGTGTTCA 1860  
Qy 621 SerSerGluGlyGlnThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640  
Db 1861 AGCAGCAGGGGCCAAACCCACACCGGAAATGGAGATGGGGAAGATTAAAGAAACCGGTGAG 1920  
Qy 641 GlyTyrSerGlyProSerLeu 647  
Db 1921 GGTATAGTGGACCTTCTCTT 1941

RESULT 3  
AX825705  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
ORIGIN  
Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:  
Best Local Similarity:  
Query Match:  
DB:  
US-10-086-464-2 (1-647) x AX825705 (1-2189)  
Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20  
Db 97 ATGTCTCGGGCGCGCTCTCCGGGAGCTGGTTCGGCTCCATCTCCACCATCAAACTCCACA 156  
Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40  
Db 157 ACCACCACTCTCTCCAGCTTCGGCTCTCTCCACCAACACCTTCTCTCTCCGCGC 216  
Qy 41 ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaProPro 60  
Db 217 CCATCCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276  
Qy 61 ProSerProProThrProSerThrProGlySerProProProProProGlnProSerPro 80  
Db 277 CCATCTCCACCACTCCATCTACCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCA 336  
Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100  
Db 337 CCGCTCCAACTACGCGCGGATCTCCACCGGACCTGTTACTCTCTCTCTCTCTCTCTCTCT 396  
Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120  
Db 397 CCACCTTTCAGTCCAGGACCACTCCCAATCTTTCGCGAAGGAGGATCTCTCTGACCT 456  
Qy 121 ProSerSerProProProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140  
Db 457 CCATCT 516  
Qy 141 GlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160  
Db 517 GGAATCGCATCGGAGGAGTCT 576  
Qy 161 LysLysLysArgArgArgAspGluAspAlaTyrValProProProProProProProPro 180  
Db 577 AAGAAGAAACACGCGAGAGACGAAAGATGCTTACTATGTGTCTCTCTCTCTCTCTCTCTCT 636  
Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200

AX825705  
Sequence 3 from Patent WO03072763.  
AX825705  
AX825705.1 GI:39751232  
Brassica napus (rape)  
Brassica napus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1  
Goring, D., Silva, N. and Haffani, Y. Z.  
Increasing plant seed production  
Patent: WO 03072763-A 3 04-SEP-2003;  
Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)  
Location/Qualifiers  
1. .2189  
/organism="Brassica napus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3708"



DB:	8	Gaps:	0
US-10-086-464-2 (1-647) x AY028699 (1-2189)			
Qy	1	MetSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr	20
Db	97	ATGTCCTCGGGCGGCTCTCGGGGACTGGTTGGCTCCATCTCCACCATCAAACTCCACA	156
Qy	21	ThrThrThrProProAlaSerAlaProProProThrThrProProSerSerProProPro	40
Db	157	ACCACCACTCCTCCTCCAGCTTCGGCTCCTCCTCCACCAACCACTTCTCTCTCCTCGCGC	216
Qy	41	ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro	60
Db	217	CCATCCACTATTCGACATCTCCTCCTCTCTTCCTCGCTCTACACCTTCTGCTCCTCCT	276
Qy	61	ProSerProProThrProSerThrProGlySerProProProLeuProGlnProSerPro	80
Db	277	CCATCTCCACCACTCCATCTACGCCGGATCTCCACCTCTCTCTCTCAGCCGCTCTCCA	336
Qy	81	ProAlaProThrThrProGlySerProProAlaProValThrProProThrProArgAsnPro	100
Db	337	CCCGCTCCAACTACGCCCGGATCTCCACCCGACCTGTACTCTCTCTCTCTCTCTCTCT	396
Qy	101	ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro	120
Db	397	CCACCTTCAGTCCCAAGGACCAACCGTCCCAATCTTCCACGGAAGGAGATCTCTCGACCT	456
Qy	121	ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValVal	140
Db	457	CCATCTCTCCTCGCCCGCTCTCTCTCTCCGACGGTTTATCAACAGGAGTGGTGGTG	516
Qy	141	GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys	160
Db	517	GGAAATCGCCATCGGAGGAGTCCGCTGCTGTGTGTAGTAGTCTGATTTGTCTCTCTGT	576
Qy	161	LysIleLysArgArgArgAspGluAspAlaTyrTyrValProProProProProPro	180
Db	577	AAGAAGAAACGACGAGAGACGAAGAAGATGCTTACTATGTTCCTCGCCACCTCTCTCT	636
Qy	181	GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnTyrArgGlnGlnAsnAla	200
Db	637	GGTCCCAAGCGGGGACCTTACGGTGGACAGCAGCAACAATGGCGGCACACAAACGCA	696
Qy	201	ThrProProSerAspHisValThrSerLeuProProProProProLysAlaProSerPro	220
Db	697	ACACCACCGTCAGATCATGTGTCGACGTCACCTACCCACCACTTAAGGCTCCATCTCCA	756
Qy	221	ProArgGlnProProProProProProProProPheMetSerSerSerGlyGlySerAsp	240
Db	757	CCACGGCAACCTCTCTCCACCTCCACCCGCTTTTCATGAGCAGCGCGGCTCCGAC	816
Qy	241	TyrSerAspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLys	260
Db	817	TACTCGGACCTGTCAGTTCTTCTCTCCACCGCTCCAGGGCTTGTGTAGGCTTCTCCAAA	876
Qy	261	SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu	280
Db	877	AGCACTTTACATACAGGAGCTAGCTAGAGCCCAATGGTTTCTCCGAGCGGCAACTTG	936
Qy	281	LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal	300
Db	937	TTAGGACAAGCGGGTTCGGTTACGTGCACAAAGGTGTGTTCCTCTAGTGGGAAAGAGTT	996
Qy	301	AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal	320
Db	997	GCTGTGAACAGCTTGAAGATTGGGAGTGGTTCAGGGAGAGAGGAGTTTCAGCAGAGTT	1056
Qy	321	GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla	340
Db	1057	GAGATCATCAGCAGAGTTTCCACACAGGCATCTGGTGTCTCTTGTGGTTATTGATCCGC	1116
Qy	341	GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuHisLeu	360

Db	1117	GGTCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAACAAACAAATCTCGAGCTTCACCTC	1176
Qy	361	HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer	380
Db	1177	CATGGCAGGAGCGCTTACAATGGAATGGAGCACCAAGATTGAAGATTGCTCTTTGGATCT	1236
Qy	381	AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle	400
Db	1237	GCTAAAGGACTTCTTATCTTCAATGAAGATTGCAATCCTAAAAATCATTCACCGGATATC	1296
Qy	401	LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu	420
Db	1297	AAGCTTCAACACATATTGATAGATTTCAGTTTGAAGTAAAGTTGCTGATTTTGGTCTT	1356
Qy	421	AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly	440
Db	1357	GCTAAGATTGCTTCTGATACAAACACGATGATCAACACGCTGATGGAACCTTTGGG	1416
Qy	441	TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer	460
Db	1417	TACTTGGCTCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTTCA	1476
Qy	461	PheGlyValValLeuLeuLeuIleThrGlyArgArgProValAspAlaAsnVal	480
Db	1477	TTTGGCGTTGCTTTTGGAGCTCATTTCTGGACGTCGACCCGTTGATGCCCAACAATGTC	1536
Qy	481	TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln	500
Db	1537	TATGTAGATGACAGCTTAGTTGACTGGGCACGACCATTTGCTTAACCGAGCATCTGAGCA	1596
Qy	501	GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet	520
Db	1597	GGAGACTTTGAGGGTTTAGCTGATGCAAGAGTAAATATGGGTATGACAGAGAGAGATG	1656
Qy	521	AlaArgMetValAlaCysAlaAlaAalaCysValArgHisSerAlaArgArgProArg	540
Db	1657	GCTCGCATGGTTGCTTGTGTCGGCTTGTGTCGCAATTCAGCTCGCCGAGACCTCGC	1716
Qy	541	MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly	560
Db	1717	ATGAGCCAGATTGTCGCTGAGAAAGAAATGTATCAGCTGTCAAGTCTTAACGAAGGG	1776
Qy	561	MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer	580
Db	1777	ATGAGCAGGTCAAAGCAATGTATACAGCTCATACGGAAGAACCCGATTATGACTCG	1836
Qy	581	SerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyr	600
Db	1837	AGCCAGTACAATGAAGACATGAAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC	1896
Qy	601	AsnAlaThrGlyGlyTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer	620
Db	1897	AACGCCACGGGTAGTACAGTAATCCGACCACTGACTATGAGCTGATGCCCTCTGGTTCA	1956
Qy	621	SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln	640
Db	1957	AGCAGCAGGCGCCAAACCAACACCGGAAATGGAGATGGGAAAGATTAAAGAACCGGTGAG	2016
Qy	641	GlyTyrSerGlyProSerLeu	647
Db	2017	GGTTATAGTGGACCTTCTCTT	2037
RESULT 5			
AX825738			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Arabidopsis thaliana (thale cress)			
Arabidopsis thaliana			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
AX825738			
Sequence 36 from Patent WO03072763.			
AX825738			
AX825738.1			
GI:39751255			
1959 bp			
DNA			
linear			
PAT 11-DEC-2003			

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

1 Goring, D., Silva, N. and Haffani, Y. Z.

Increasing plant seed production

Patent: WO 03072763-A 36 04-SEP-2003;

JOURNAL Daphne (CA); Silva, Nancy (CA); Haffani, Yosr, Z. (CA)

Goring, Location/Qualifiers

## FEATURES

1. 1959

source /organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3702"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,23e-63 Length: 1959  
Score: 2936.50 Matches: 574  
Percent Similarity: 90.61% Conservative: 24  
Best Local Similarity: 86.97% Mismatches: 41  
Query Match: 85.04% Indels: 21  
DB: 6 Gaps: 12

US-10-086-464-2 (1-647) x AX825738 (1-1959)

Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17  
Db 1 ATGTCACAGCGCGTCTCCAGGCACTACTCCATCACCCTCCACCGTCTCTCCACCA 60  
Qy 18 AsnSerThrThrThrProProProAla-----SerAlaProProThrThrPro 35  
Db 61 AACTCGCAACACCACTCTCTCCAGCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
Qy 36 SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThrThr 54  
Db 121 TCTTCT 180  
Qy 55 ThrProSerAlaProProProSerProProThrProProThrProGlySerProPro 74  
Db 181 CCT 228  
Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94  
Db 229 CTTCCTCAACCTTCCCTCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 285  
Qy 95 ProProThr-----ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113  
Db 286 CCCTCAACCTCTGAAGCCCTCTCACTCTCTCAACCAAGGACCAACACCTCTCTCTCA 342  
Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131  
Db 343 ---GGATCTACTCTAGAACTCCATCAACACTAAACCGTCCGCGCTCTGATCTCTCC 399  
Qy 132 AspGlyLeuSerThrGlyValValValGlyLeuAlaLeuGlyValAlaLeuLeuVal 151  
Db 400 GATCGATGTTCTACCGGAGTTGTGTAGGAATCGCCATTTGGAGGAGTCTGATTTCTG 459  
Qy 152 IleValThrLeuIleCysLeuLeuCysIleCysIleCysIleCysIleCysIleCys 169  
Db 460 ATACTGACTCTGATTTGTTCTCTCTGTAAGAAAGAAACGAAGAAACGACGATGAA 519  
Qy 170 AspAlaTyrTyrValProProProProProGlyProProGlyProProGlyProProGly 189  
Db 520 GCTGCTTACTATGTTCT 579  
Qy 190 GlyGlnGlnGlnInTrpArgGlnGlnAsnAlaThrProProSerAsp-----HisValVal 208  
Db 580 GGTCAACAACTATTGCAACAAACAAACCGCTCAGCGCCGTCAGATATCATGTAGTG 639  
Qy 209 ThrSerLeuProProProProProProProProProProProProProProProProPro 228  
Db 640 ACGTCATTG-----CCACCACTTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 696  
Qy 229 ProProPro-----PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247

Db 697 CCACCACGACCATTCATGAGTAGCAGTGGTGTCTGACTATTTCGGATCTTCCGGTCTCT 756  
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerIysSerThrPheThrTyGluGlu 267  
Db 757 CTTCCACCATCTCCAGGGCTTGTGTAGGCTTTTCTAAAAGCACTTTTACATATAGAGAG 816  
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287  
Db 817 TTGTGAGAGACTACTAATGCTTCTCTGAGGCTAATTTGTAGGCAAGGAGGGTGTGGT 876  
Qy 288 TyrValHisIysGlyValLeuProSerGlyIysGluValAlaValIysGlnLeuIysVal 307  
Db 877 TATGTGCATAAAGGTATATTGCCCTAGTGGGAAAGAGTTGCTGTGAACACAGTTCGAA 936  
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327  
Db 937 GGTAGTGGTCAGGAGAGAGAGAGTTCAGGCTGAGGTTGAGATCATTTAGCAGAGTTT 996  
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaIysArgLeuVal 347  
Db 997 CACAGGCAATTTGGTTTCTCTTATTGGTTATTGTATGGCCGGTGTTCAAAGATTACT 1056  
Qy 348 TyrGluPheValProAsnAsnAsnLeuGluLeuHisGlyGlyGlyGlyGlyGlyGly 367  
Db 1057 TATGAGTTTGTTCACAAACAATCTTGTAGTTTCCCTCCATGCTTAAGGAGCGCTACG 1116  
Qy 368 MetGluTrpSerThrArgLeuIysIleAlaLeuGlySerAlaIysGlyLeuSerTyrLeu 387  
Db 1117 ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGGATCTGCTAAAGGACTTTTCTAT 1176  
Qy 388 HisGluAspCysAsnProIysIleIleHisArgAspIleIysAlaSerAsnIleLeu 407  
Db 1177 CATGAAGATTGCAATCCGAAATCATTCACCGTGATTAAGGCGTCAACATATTGATT 1236  
Qy 408 AspPheIysPheGluAlaIysValAlaAspPheGlyLeuAlaIysIleAlaSerAspThr 427  
Db 1237 GATTTCAAATTGAAGCTAAGTGTGCTGCTTGGTCTTGCACAGATTGCTTCTGTATACA 1296  
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGlyTyrAla 447  
Db 1297 AACACTCATGTATCTACACGCTGATGGAACCTTTTGGGTATTTGGCTCCGGAATATG 1356  
Qy 448 AlaserGlyIysLeuThrGluIysSerAspValPheSerPheGlyValValLeuLeuGlu 467  
Db 1357 GCAAGTGAAGAGCTCACAGAAAGTCTGAGTTTCTCATTTGGCGTTGTACTTTGGAA 1416  
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487  
Db 1417 CTTATTACTGGAGGCGCCCTGTTGATGCGAACATGCTTATGTATAGATGACAGCTTAGTT 1476  
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507  
Db 1477 GACTGGGCACGACCATTTGCTTAAACCGGATCTGAGGAGGAGATTTTGGAGGTTTGGCT 1536  
Qy 508 AspAlaIysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527  
Db 1537 GATTCAAGATGGGTAAATGATGATGACAGAGAGAGATGCTCGCATGTTGTTGGCT 1596  
Qy 528 AlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGlnIleValArgAla 547  
Db 1597 GCGCTTGTCTGCCCATTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTAGCGGCG 1656  
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567  
Db 1657 TTAGAAGGAAATGTAATCGCTGCTGATCTTAAACGAAGGAGTGAACCGGCTCACAGAAC 1716  
Qy 568 ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587  
Db 1717 GTATACAGCTCATATGAGGAGGAGCAGACTATGACACGAGCCCATCAACAGCAGCATG 1776  
Qy 588 LysIysPheArgLysMetAlaLeuGlyThrGlnGlyTyrAsnAlaThrGlyGlyTyrSer 607



190 GlyGlnGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208  
 Db GGTCAACAACAGTATTGGCAACAACAACCGGTACAGCCGCTCAGATAATCATGTAGTG 639  
 209 ThrSerLeuProProProProGlyAlaProSerProProArgGlnProProProProPro 228  
 Db ACAGTCATTTG---CCACACACCTTAACCTCCATCTCCACCAACCAACCTCTCCGCCACCT 696  
 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247  
 Db CCACACACAGCATTCATGAGTACGAGTGGTGGTCTCGACTATTCGATCTCCGGTCTTT 756  
 248 ProProProSerProGlyLeuValLeuGlyPheSerLeuSerThrPheThrTyrGluGlu 267  
 Db CCTCCACCATCTCCAGGGCTTGTGTAGGCTTTTCTAAAGACATCTTCACTTATGAGGAG 816  
 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287  
 Db TTGTCGAGAGCTACTAATATGGCTTCTCTGAGGCTAATTTGTTAGGACAGGAGGGTTTGGT 876  
 288 TyrValHisGlyValLeuProSerGlyValValAlaValValLeuValValValVal 307  
 Db TATGTGATTAAGTATATTCCTAGTGGGAAAGAGTTGCTGTGAACAGTTGAAAGCT 936  
 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuLeuSerArgValHis 327  
 Db GGTAGTGGTCAGGAGAGAGAGAGTTCAGGCTCAGGTTGAGATCATTAGCAGAGTTTCAT 996  
 328 HisArgHisLeuValSerLeuValGlyTyrCysLeuAlaGlyAlaLysArgLeuLeuVal 347  
 Db CACAGGATTTGGTTCTCTTATTGGTTATTGTATGGCCGGTGTTCAAAGATTAATTCTTGT 1056  
 348 TyrGluPheValProAsnAsnAsnLeuGluLeuHisGlyGlyGlyGlyGlyArgProThr 367  
 Db TATGAGTTTGTCCAAACAACATCTTGAGTTTCACTCTCATGTATGAAGGACGTCCTACG 1116  
 368 MetGluTrpSerThrArgLeuLysLeuAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387  
 Db ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGATCTGCTAAAGGACITTCATATCTT 1176  
 388 HisGluAspCysAsnProLysLeuLeuHisArgAspLeuLysAlaSerAsnLeuLeu 407  
 Db CATGAAGATTGCAATCCGAAATCATTCACCGTGATTAAGGCGTCAAAACATATTGATT 1236  
 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysLeuAlaSerAspThr 427  
 Db GATTTCAAATTTGAAGTAAAGTTGCTGACTTTGGTCTTGCCAAAGATTGCTTCTGATACA 1296  
 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447  
 Db AACACTCATGTACTACACGGTGATGGAAACCTTTGGGTATTGGTCCGGAATATGCT 1356  
 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467  
 Db GCAAGTGGAAAGCTCACAGAAAAGCTGACGTTTCTCAITTTGGCGTTGTACTTTTGAA 1416  
 468 LeuLeuThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487  
 Db CTATTATTCTGGAGGCCCTGTTGATGGCAACAATGCTCATGATGATGACAGCTTAGTT 1476  
 488 AspTriaAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507  
 Db GACTGGGACGACCATTTGCTTACCGAGCATCTGAGGAAGAGATTTTGGGGTTGGCT 1536  
 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527  
 Db GATTTCAAAGATGGGTAATGATGATGACAGAGAGAGATGCTCGCATGGTTGCTTGGCT 1596  
 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnLeuValArgAla 547  
 Db GCAGCTTGTGTGCCATTGAGTCTCGCCGAGAGCTTCGATGAGCCAGATAGTACGGGCG 1656  
 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567

1657 TTAGAAGAAATGATCGCTGTCTGATCTTAACGAAGGATGAGACCGGGTCACAGCAAC 1716  
 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587  
 1717 GTATACAGCTCATATGAGGAAGCACAGATATGACAGGACCAATACACAGCAGCATG 1776  
 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607  
 1777 ATAAAGTTTAGGAAATGGCTCTTGGAACTCAAGAAATACGGCACAAACGGGAGTACAGT 1836  
 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627  
 1837 AATCCAAACCAAGTACTACGAGCTGTACCGCTCTGTTCAAGCAGTGAAGGTCAAGCCACA 1896  
 628 ArgGluMetGluMetGlyLysLeuLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647  
 1897 CGAGAAATGAGATGGGAAAGATCAAGAAACCGGTCAAGGTTATAGTGGACCTCTCTT 1956  
 RESULT 7  
 LOCUS AY128792 2098 bp mRNA linear PLN 07-AUG-2002  
 DEFINITION Arabidopsis thaliana protein kinase-like protein mRNA, complete cds.  
 ACCESSION AY128792  
 VERSION AY128792.1 GI:22136227  
 KEYWORDS FLJ CDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 2098)  
 REFERENCE Tripp,M., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
 AUTHORS Direct Submission  
 TITLE Submitted (01-JUL-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA  
 JOURNAL  
 COMMENT e-mail for correspondence: arab@sequence.stanford.edu  
 The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.  
 The Salk, Stanford, POGC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Tripp,M., Nguyen,M., Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Tripp,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.  
 FEATURES  
 Location/Qualifiers  
 1..2098  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /clone="U16568"  
 /ecotype="Columbia"  
 /note="This clone is in pENTR/SD-dTopo This is a cloned PCR product using RIKEN clone RAFL09-06-A17 (AY059901) as a template"  
 1..1959  
 /codon\_start=1  
 CDS









Qy 568 ValTyrSerSerTyrGlySerThrAspTyrAspSerSerSerGlnTyrAenGluAspMet 587  
 Db 1748 GTATACAGCTCATATGGAGGAAGCACAGACTATACACAGAGCCCAATCAACAGCAGCATG 1807  
 Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGlnTyrAenAlaThrGlyGluTyrSer 607  
 Db 1808 ATAAAGTTAGAAAATGGCTCTTGGAACTCAAGNAATACGCACACCGGGAGTACAGT 1867  
 Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627  
 Db 1868 AATCCAAACCACTGACTACGAGCTACCGCTCTGGTTCAAGCACTGAAGGTCAAGCCACA 1927  
 Qy 628 ArgGluMetGluMetGlyLysLeuLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647  
 Db 1928 CGAGAAATGGAGATGGGAAGATTAAAGAAACCGGTCAAGGTATTATAGTGGACCTCTCTT 1987

RESULT 10  
 AV093065  
 LOCUS  
 DEFINITION Arabidopsis thaliana unknown protein (At3g24550) mRNA, complete cds.  
 ACCESSION AY093065  
 VERSION AY093065.1 GI:20260331  
 KEYWORDS FLI CDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 1 (bases 1 to 2190)  
 Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Direct Submission  
 Submitted (27-MAR-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA  
 e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA': Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Southwick,A., Nguyen,M., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES  
 source  
 Location/Qualifiers  
 1..2190  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 /clone="RAFL07-12-E08 (R17522)"  
 /ecotype="Columbia"  
 /note="This clone is in pBluescript"

gene  
 1..2190  
 /gene="At3g24550"  
 30..1988  
 /gene="At3g24550"  
 /codon\_start=1  
 /product="unknown protein"

/protein\_id="AAAM13064.1"  
 /db\_xref="GI:20260332"  
 /translation="MSTAPSGCTTPSPSPPTNSTTTTPPPAASPPPTTTTSSPP  
 PSSTNSTSPSPSPPLPSPSPGSLTPPLPSPSAPITPPSPSPPTTSPSPSP  
 PSNQGPNTPTSGTPTPTSPNTKSPSPSDSLGTVVGVGAIAGGVAILVILTLCL  
 LCKKRRRRHDDAAAYVPPPPSPKAGPGYQQYQQQNASRSDNVHTSLPP  
 PKPSPRPKPPPPPPAPPPSSGSDYSDLPVLPSPSGLVGFSKSTFYELSLRA  
 TNGFSANLLGGGFGVHKGILPSSKEVAVKOLKAGSGGGEREFOAEVEIISRVHR  
 HLVLGYCNAGVQRLLVYEFVNNNLEFHLHGKPTMWSITRLKALASGSLSVL  
 HEDCNPKIHRDKIKSNLILDFPEAKVADFLAKTASDINTHVRWMTGTFGLAE  
 YAAAGKLTEDKSFVSGVLLLELTIGRRPVANNVYVDDSLVWARFLNARSEGEF  
 EPLASKMGNEYDEEMARMVCAACAAKHSARRPRMSQIVRALEGNVSLDNEG  
 RGHNSVYSYGASTDYDTQYNDMLKFRKMALGTQYGTGTGEYSNPTS DYGLYPSG  
 SSSEGOATREMENMKIKKTCQYSGPSL"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.33e-63 Length: 2190  
 Score: 2936.50 Matches: 574  
 Percent Similarity: 90.61% Conservative: 24  
 Best Local Similarity: 86.97% Mismatches: 41  
 Query Match: 85.04% Indels: 21  
 DB: 8 Gaps: 12  
 US-10-086-464-2 (1-647) x AY093065 (1-2190)  
 Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17  
 Db 30 ATGTCCACAGCGCGTCTCCAGGCACTACTCCATCACCACATCTCCACCGTCTCTCCACCA 89  
 Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrThrPro 35  
 Db 90 AACTCGAACACCACTCTCTCTCCAGACGCTTCTCTCTCTCCACCACTCTCTCTCTCTCT 149  
 Qy 36 SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThrThrThr 54  
 Db 150 TCTTCT 209  
 Qy 55 ThrProSerAlaProProProSerProProThrProThrProThrProGlySerProProPro 74  
 Db 210 CCTCCTTCT 257  
 Qy 75 LeuProGlnProSerProAlaProThrThrProGlySerProProAlaProValThr 94  
 Db 258 CTCTCTCAACCTTCCCGCTCCGCTCCCATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 314  
 Qy 95 ProProThr-----ArgAsnProProProSerValProGlyProProProSerAsnProSerArg 113  
 Db 315 CCCTCAAAACCTCGAAGCCCTCCATCTCTTAACCAAGGACCAACCAACACATCTCTCTCA--- 371  
 Qy 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131  
 Db 372 ---GGATCTACTCTAGAACTCCCAACAACTAAACCGTCCGCGCGCTCTGATCTCTCTCC 428  
 Qy 132 AspGlyLeuSerThrGlyValValGlyValAlaIleGlyGlyValAlaLeuVal 151  
 Db 429 GATGGATTGTTACCGAGTTGTGTAGTAATCCCATTCGAGGAGTTCGATTTCTTGT 488  
 Qy 152 IleValThrLeuIleCysLeuLeuCysLysLysLysArgArg-----AspGluGlu 169  
 Db 489 ATACTGACTCTGATTTGTCTTCT 548  
 Qy 170 AspAlaTyrThrValProProProProGlyProGlyProlysAlaGlyGlyProTyrGly 189  
 Db 549 GCTGCTTACTATGTCT 608  
 Qy 190 GlyGlnGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208  
 Db 609 GGTCAACACAGTATTGGCAACAACAAACCGCTCGCGCGCTCTAGATAATCATGATG 668  
 Qy 209 ThrSerLeuProProProProGlyAlaProSerProArgGlnProProProProProProPro 228  
 Db 669 ACGTCATTG---CCACCACCTTAAGCTCCATCTCCACCCAGCAAAACCTCTCTCTCCGCACT 725

Qy 229 ProProPro----PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247  
Db 726 CCACCACCAGCATTCATGAGTAGCAGTGGTCTGCTGACTATTCGGATCTTCGGTCTCT 785  
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyGluGlu 267  
Db 786 CTTCCACCATTCTCAGGGCTTGTGTAGGCTTTCTTAAAGACACTTTTCACTTATGAGGAG 845  
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287  
Db 846 TTGTCGAGAGCTACTAATGGCTTCTCTGAGGCTTAATTTGTAGGACAGGAGGGTTGGT 905  
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307  
Db 906 TATGTGCATAAAGTATATTCCTAGTGGGAAAGAGTTGCTGTGMAACAGTTCGAAGCT 965  
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHis 327  
Db 966 GGTAGTGTCTCAGGGAGAGAGAGATTTTCAGGCTGAGGTTGAGATCATTTAGCAGAGTTTCA 1025  
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysLeuLeuAlaGlyAlaLysArgLeuVal 347  
Db 1026 CACAGGCAATTTGGTTCCTTATTTGGTTATTTGATGGCCGGTGTTCAAAGATTACTTTGTC 1085  
Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367  
Db 1086 TATGAGTTTGTTCACAAACAATCTTGAGTTTCACCTCCATGTTGTAAGGAGCGCTACG 1145  
Qy 368 MetGluTrpSerThrArgLeuLysLeuAlaLeuGlySerAlaLysGlyLeuSerTyLeu 387  
Db 1146 ATGGAATGGAGTACTAGATTGATGATTCGATCTGCTGATGCTGCTGCTGCTGCTGCTGCT 1205  
Qy 388 HisGluAspCysAsnProLysLeuLeuHisArgAspLeuLysAlaSerAsnLeuLeu 407  
Db 1206 CATGAAGATTTCGAATCCGAATCATTCACCGTGATATTAGGGCTCAACATATTGATT 1265  
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysLeuLeuSerAspThr 427  
Db 1266 GATTTCAAATTTGAAGCTAAGTTGCTGACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1325  
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyAla 447  
Db 1326 AACACTCATGTATCTACACGGCTGATGGAACTTTGGGTATTTGGCTCCGGAATATGCT 1385  
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467  
Db 1386 GCAAGTGGAAAGCTCAGAAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGAA 1445  
Qy 468 LeuLeuThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487  
Db 1446 CTTATTACTGGAGGCGCCCTGTTGATGCGGACCAATGCTATGTATGATGACAGCTTAGTT 1505  
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507  
Db 1506 GACTGGGCACGACCATTTGCTTAACCGAGCATCTCAGGAAGGAGATTTTGAGGGTTGGCT 1565  
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAla 527  
Db 1566 GATTCAAGATGGGTAATGATGATGACAGAGAGAGATGGCTCCCATGGTTGCTTGGCT 1625  
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnLeuValArgAla 547  
Db 1626 CGCGCTGTGTTCGCCATTAGCTCGCGGACAGCTCGCATGACCGATAGTAGTACGGCG 1685  
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567  
Db 1686 TTAGAAGAAATGATCGCTGCTGCTGCTTAAACGAAGGATGAGACCGGGTCAACGCAAC 1745  
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587  
Db 1746 GTATACAGCTCATATGAGGAAGCACAGACTATGACACGAGCCCAATCAACAGCAGCATG 1805

Qy 588 LysLysPheArgLysMetalAlaLeuGlyThrGlnGlyTyrAsnAlaThrGlyGlyTyrSer 607  
Db 1806 ATAAAGTTTAGGAAAATGGCTCTTGGAACTCAAGAAATACGCACACCGCGAGTACAGT 1865  
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnGlyGlnThr 627  
Db 1866 AATCCAAACCACTGACTACGAGCTGTACCCGCTCTGGTTCAAGCAGTGAAGGTCACAGCCACA 1925  
Qy 628 ArgGluMetGluMetGlyLysLeuLeuLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647  
Db 1926 CGAGAAATGAGATGGGAAAGATTAAAGAAACCGTCAAGTTATAGTGGACCTCTCTT 1985  
RESULT 11  
BT008409 1959 bp mRNA linear PLN 15-MAY-2003  
LOCUS Arabidopsis thaliana At3g24600 gene, complete cds.  
DEFINITION Arabidopsis thaliana  
ACCESSION BT008409  
VERSION BT008409.1 GI:30725491  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1 (bases 1 to 1959)  
REFERENCE  
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,  
Dale,J.M., Hayashizaki,Y., Heuan,V.W., Ishida,J., Jones,T.,  
Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,  
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,  
Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,  
Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,  
Theologis,A. and Ecker,J.R.  
TITLE Arabidopsis ORF clones  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1959)  
AUTHORS Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Bowser,L., Carninci,P.,  
Dale,J.M., Hayashizaki,Y., Heuan,V.W., Ishida,J., Jones,T.,  
Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M.,  
Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L.,  
Sakurai,T., Satou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C.,  
Wu,H.C., Yamada,K., Yu,G., Yuan,S., Shinozaki,K., Davis,R.W.,  
Theologis,A. and Ecker,J.R.  
TITLE Direct Submission  
JOURNAL Submitted (15-MAY-2003) Salk Institute Genomic Analysis Laboratory  
(SiGnAL), Plant Biology Laboratory, The Salk Institute for  
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,  
USA  
COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,  
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,  
Hayashizaki,Y. and Shinozaki,K.  
The Salk, Stanford, PGECC (SSP) Consortium members constructed and  
sequenced the PUNI (ORF) clones using the RAFL cDNAs: Cheuk,R.,  
Chen,H., Kim,C.J., Shinn,P., Bowser,L., Chan,M.M., Chang,C.M.,  
Dale,J.M., Heuan,V.W., Jones,T., Karlin-Neumann,G., Lam,B.,  
Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J.,  
Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,  
Yamada,K., Yu,G., Yuan,S., Davis,R.W., Theologis,A., and  
Ecker,J.R.  
Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to  
this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk)  
contributed equally to this work as PIs.  
FEATURES  
source  
1. 1959  
Location/Qualifiers  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="U15366"  
/ecotype="Columbia"



```

Qy 568 ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1717 GTATACAGCTCATATGAGGAGGACACAGACTATGACAGCGCAATACAGCAGCATG 1776

Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1777 ATAAAGTTTAAAGAAATGGCTCTTGGAACTCAAGAAATACGGCACACCGCGGAGTACAGT 1836

Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThr 627
Db 1837 AATCAACCACTGACTACGGACTGTACCGTCTGTTCAAGCAGTGAAGTCAAGCCACA 1896

Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647
Db 1897 CGAGAAATGGAGATGGGAAAGATTAGAAACCGGTCAAGGTTATAGTGGACCTCTCTT 1956

RESULT 12
AF370509 2257 bp mRNA linear PLN 30-APR-2001
LOCUS Arabidopsis thaliana protein kinase-like protein (MOB24.13) mRNA,
DEFINITION complete cds.
ACCESSION AF370509.1 GI:13877616
VERSION FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 2257)
Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,
Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (17-APR-2001) DNA Sequencing and Technology Center,
Stanford University, 955 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Southwick,A.,
Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J.,
Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.
(SSP/Stanford) contributed equally to this work as PIs.

FEATURES
Location/Qualifiers
1..2257
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="III"
/clone="RAFL11-07-N23"
/ecotype="Columbia"
/note="This clone is in pBluescript"
1..2257
/gene="MOB24.13"
107..2065
/gene="MOB24.13"
/codon_start=1

gene
CDS

```

```

/product="protein kinase-like protein"
/protein_id="AAK43886.1"
/db_xref="GI:13877617"
/translation="MSTAPSPGTTTPSPSPPTNSPTTTPPAASPPPTTTTPSSPP
PSSTNGSTSPSPSPPLPSLPSPSPSGSLTPPIQPSPSPADIPSPSPPTPSNPSP
PSNQPPNTPSGSTPRTNSNTPSPSPSSDGLSTGVVGIAGVAILVILTLICL
LCKKRRRRHDDAAAYVPPPPSPGKAGPYGGQQYQQQNNRPSDNRHVTSLLP
PKPPSPRPKPPPPPPAFMSSSGSDYDLVLPSPSPGLVLFKFSKSTFTYELSRA
TNGFSANLGGQGGFGVHKGLIPSGKEVAVKOLKAGSGGGERFOAEVEIISRVHHR
HLVSLICYMAGVQRLVVEFVNNLEHLHGKGRPTMESTRLKIALGSAKGLSYL
HEDCNPKIIRHDIKASNILIDPFEAKVADPGLAKIADSTNTHVSTRVMTGTFGLAPE
YASGKUTEKSDVFSFGVLELITGRPDANNVYVDDSLVWARLPLLRASSEGDF
EGLADSKMGNEYDREEMARVACAAACHSARRRPMRSQIVLALGNVSLSDLEMG
RPGHSNVSYSGGSDYDTSYNDMDIKFRKMAIGTQEYGTTCGETSNPTSDYGLYPSG
SSSEGGQATREMENGKIKTKQYSGPSL"

ORIGIN
Alignment Scores: 1.52e-63 Length: 2257
Pred. No.: 2934.50 Matches: 573
Score: 2934.50
Percent Similarity: 90.61% Conservative: 25
Best Local Similarity: 86.82% Mismatches: 41
Query Match: 84.98% Indels: 21
DB: 8 Gaps: 12

US-10-086-464-2 (1-647) x AF370509 (1-2257)
Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
Db 107 ATGTCCACAGCGCGCTCCAGGACCTATCCATCCATCTCCACGCTCTCTCCACCA 166
Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db 167 AATCGACACACCACTCTCTCCAGAGCGCTCTCTCTCTCTCCACCACTCTCT 226
Qy 36 SerSerProProPro-----ProSerThrThrThrThrThrThrThrThrThr 54
Db 227 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 286
Qy 55 ThrProSerAlaProProProSerProProThrProThrProGlySerProPro 74
Db 287 CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 334
Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
Db 335 ATTCTCTCAACTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 391
Qy 95 ProProThr-----ArgAsnProProProSerValProGlyProProSerAsnProSerArg 113
Db 392 CCTCAAAACCTCGAAGCCTCCATCTCTTAACCAAGGACCAACCAACACTCTCTCA 448
Qy 114 GluGlyGlySerProProProProSerSer-----ProSerProProSerProSer 131
Db 449 ---GGATCTACTCTCTGAATCTCAACACTCAAAACCGTCCGCGCGCTCTGATCTTCC 505
Qy 132 AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyGlyValAlaLeuVal 151
Db 506 GATGGATTGTCTACCGGAGTTGGTAGGATGCCATTCGAGGAGTCTGATCTTCTGTT 565
Qy 152 IleValThrIleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db 566 ATACTGACTCTGATTTGTCTCTCTCTGTAAGAAACGAAGAAAGACACGACGATGAA 625
Qy 170 AspAlaTyrTyrValProProProProProProGlyProGlyProGlyProGly 189
Db 626 GCTGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 685
Qy 190 GlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 208
Db 686 GGTCAACACAGTATTGGCCACAAACCAACCGTTCACGCGCGCTCAGATCATCTAGTG 745
Qy 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProPro 228
Db 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProPro

```

Db 746 ACGTCATTG---CCACCACCTAAGCCTCCATCTCCACCACGAAACCTCTCCGCCACCT 802  
Qy 229 ProProPro----PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247  
Db 803 CCACCACGAGCATTCATGAGTAGCAGTGGTGGTCTGACTATTCCGATCTCTCCGGTCTTT 862  
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267  
Db 863 CCTCCACATCTCCAGGGCTTGTGTAGGCTTTCTTAAAGCACTTTACACTTATGAGGAG 922  
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287  
Db 923 TTGTGAGAGCTACTAATGGCTTCTCTGAGGCTAATTTGTAGCAAGAGGGTGTGGT 982  
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307  
Db 983 TATGTGCATAAAGTATATAGCTAGTGGGAAAGAGTTGCTGTGAAACAGTGTGAAAGCT 1042  
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLysSerArgValHis 327  
Db 1043 GGTAGTGGTCAAGGAGAGAGAGTGTTCAGGCTGAGGTTGAGATCATTTAGCAGAGTTTCAT 1102  
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347  
Db 1103 CACAGGCAATTTGGTTTCTCTTATTTGGTATTGTATGATGCCGGTGTCAAAGATTTACTTGT 1162  
Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlyArgProThr 367  
Db 1163 TATGAGTTGTTCACAAACAAACAACTTGTGAGTTTCACTCCATGTTGTAAGGAGCGCTACG 1222  
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387  
Db 1223 ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGATCTGCTAAAGGACTTTTCATATCTT 1282  
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407  
Db 1283 CATGAGATTGCATTCGAAATCATTCACCGTATATTAAAGGCTCAACATTTGATTT 1342  
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427  
Db 1343 GATTTCAAATTTGAAGCTAAGGTTGCTGACTTTGCTTGTGCAAGATTTGCTTCTGATACA 1402  
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447  
Db 1403 AACACTCATGTATCTACACGGCTGATGGAAACCTTTGGGTATTTGGCTCCGGAATATGCT 1462  
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467  
Db 1463 GCAAGTGGAAAGCTCACAGAAAGCTGACGTTTCTCATTTGGCGTTGTACTTTTGAA 1522  
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487  
Db 1523 CTTATTACTGGAGGCGGCTGTTGATGCGAACAATGCTATGTATGATGACAGCTTAGTT 1582  
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507  
Db 1583 GACTGGGCACACCATTCGTTTAAACCGACATCTGAGGAAGAGATTTTGAAGGGTTTGGCT 1642  
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527  
Db 1643 GATTCAAAGATGGGTATGAGTATGACAGAGAGAGATGGCTCCGATGGTTGCTTGGCT 1702  
Qy 528 AlalaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547  
Db 1703 GCGGCTTGTGTCCGATTCAGTCCGCGACAGCTCGCATGAGCCAGATAGTACGGGCG 1762  
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567  
Db 1763 TTGAAGGAAATGATTCGCTCTGATCTTAAAGAGGGATGAGACCGGGTCACAGCAAC 1822  
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587  
Db 1823 GTATACAGCTCATATGGAGGAGGACACAGACTATGACACGAGCCCAATACACAGCACATG 1882

Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607  
Db 1883 ATAAAGTTTAAAGAAATGGCTTTGGAACTCAAGAATACGGCACACCGGCGAGTACAGT 1942  
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGluGlnThrThr 627  
Db 1943 AATCCAAACAGTAGTACGAGCTGTACCGCTCTGGTTCAAGCAGTGAAGGTCAGGCCACA 2002  
Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647  
Db 2003 CGAGAAATGAGATGGGAAAGATTAGAAAACCGGTCAAGGTTATAGTGGACCTCTCTT 2062

## RESULT 13

AY089024 LOCUS 2324 bp mRNA linear PLN 14-APR-2003  
DEFINITION Arabidopsis thaliana clone 17909 mRNA, complete sequence.

ACCESSION AY089024

VERSION AY089024.1 GI:21407798

KEYWORDS FLU CDNA.

SOURCE Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 2324)

Haas,B.J., Volkovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,

Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.

Full-length messenger RNA sequences greatly improve genome

annotation

Genome Biol. 3 (6), RESEARCH0029 (2002)

22088475

PUBMED 12093376

REFERENCE 2 (bases 1 to 2324)

Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and

Feldmann,K.

Full-length cDNA from Arabidopsis thaliana

Unpublished

3 (bases 1 to 2324)

Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and

Feldmann,K.

Direct Submission

Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,

Malibu, CA 90265, USA

This clone sequence is one of 5,000 Ceres full-length cDNAs made of

available to TIGR and Genbank. The following quality assessment of

this set was done by comparison with known proteins: two percent of

the clones are estimated to be 5'-truncated; less than one percent

are 3'-truncated; approximately two percent represent alternative

splice variants, including unspliced introns and spliced exons; one

percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be

5'-truncated if it lacks the translation initiation start (ATG). A

sequence is considered to be 3'-truncated if it lacks the

C-terminal end of the encoded protein. Please note that these cDNA

sequences are derived from the Ws or Laer ecotypes and therefore

may contain polymorphisms when compared to sequences from Col-0.

Genet carried out the library production and sequencing of the

full-length clones. Ceres, Inc. carried out the clustering of the

5' sequences, selection of clones, and sequence assembly.

## FEATURES

source

1. .2324

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/clones="17909"

## ORIGIN

Alignment Scores:

Pred. No.: 2,45e-62 Length: 2324

Score: 2885.00 Matches: 570

Percent Similarity: 90.30% Conservatave: 26

Best Local Similarity: 86.36% Mismatches: 42

Query Match:	83.55%	Indels:	23
DB:	8	Gaps:	13
US-10-086-464-2 (1-647) x AY089024 (1-2324)			
Qy	1	MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer	17
Db	105	ATGTCACAGCGCGTCTCCAGGACACTACTCAYCACCATCTCACCGTCTCTCCGCCACA	164
Qy	18	AsnSerThrThrThrProProAla-----SerAlaProProProThrThrPro	35
Db	165	AACTCGACAACACCACTCTCTCCAGCAGCTCTTCTCTCTCTCTCCACCAACCACTCT	224
Qy	36	SerSerProProPro-----ProSerThrThrProThrSerProProSerSerArgSer	54
Db	225	TCCTTCTCTCTCTCGTCCGCGCTCAACTAAATTAACACCTCTCTCTCTCTCTCTCTCT	284
Qy	55	ThrProSerAlaProProProSerProProThrProThrProGlySerProProPro	74
Db	285	CCTCTCTCTCTCTCTCTCAATCTCTCTCTGAGCTTTAACT-----CCTCCT	332
Qy	75	LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr	94
Db	333	CTTCTCTCAACCTTCTCCATCCGCTCCCATCACTCCT---TCCTCCGCGCTCTCCCAACCA	389
Qy	95	ProProThr---ArgAsnProProProSerValProGlyProProSerProProSerArg	113
Db	390	CCCTCAAAACCTCGAAGCCCTCCATCTCTCAACCAAGGACCAACAAACACTCCCTCA	446
Qy	114	GlucGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer	131
Db	447	---GGATCTACTCTCTAGNACTCCATCAACCGCTAAACCGTCGCCACCGTCTGATCTCTCC	503
Qy	132	AspGlyLeuSerThrGlyValValValGlyIleAlaIleGlyValAlaLeuVal	151
Db	504	GATGGATTGTCTACCGGAGTTGTGTAGTAATCGCCATTGGAGGAGTCGCTATTCTTGT	563
Qy	152	IleValThrLeuIleCysLeuLeuCysIlyslYslYsArgArgArg-----AspGluGlu	169
Db	564	ATACTGACTCTGATTGTCTTCTGTAAAGAAAGAAAGAAAGAACACGACGACGATGA	623
Qy	170	AspAlaIleValProProProProProGlyProLysAlaGlyGlyProTyrGly	189
Db	624	GCTGCTTACTATGTT---CCTCTCTCTCTCTCTGTCCTCCCAAGCTGGAGGACCTTACCGT	680
Qy	190	GlyGlnGlnGlnInTrpArgGlnGlnAsnAlaThrProProSerAsp---HisValVal	208
Db	681	GGTCAACAACAGTATTGGCAACAACAAACCGCTCACGGCGCTCAGATAATCATGTAGTGT	740
Qy	209	ThrSerLeuProProProLysAlaProSerProProArgGlnProProProProPro	228
Db	741	ACGTCATG---CCACCACCTTAAGCTCTCATCTCCACCACCAAAACCTCTCTCCGCACT	797
Qy	229	ProProPro---PheMetSerSerSerGlySerAspTyrSerAspArgProValLeu	247
Db	798	CCACCACGACGATTCATGAGTAGCAGTGGTGTCTGACTATTCTCGATCTTCCGCTTCTT	857
Qy	248	ProProProSerProGlyLeuValLeuGlyPheSerIysSerThrPheThrTyGluGlu	267
Db	858	CCTCCACCATCTCAGGGCTTGTGTAGGCTTTTCAAAAAGCACTTTCACCTATTAGGAG	917
Qy	268	LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly	287
Db	918	TTGTCAAGAGCTACTAATGGCTTCTCTGAGGCTTAATTTGTAGGACAAGAGGGGTTTGGT	977
Qy	288	TyrValHisIysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal	307
Db	978	TATGTGATAAAGGTATATTTCCTAGTGGGAAGAAAGTTGCTGTGTAACACGTTCAAAGCT	1037
Qy	308	GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis	327
Db	1038	GGGAGTGTCTCAGGAGAGAGAGTTCAGCGTAGGTTGAGATCATTTAGCAGAGTTTCAT	1097

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

AUTHORS  
TITLE

1. (sites)  
Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.  
Structural analysis of Arabidopsis thaliana chromosome 3. II.  
Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC  
and BAC clones

DNA Res. 7 (3), 217-221 (2000)

## JOURNAL

## MEDLINE

## PUBMED

10907853

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

2. (bases 1 to 79706)  
Kaneko,T., Kato,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.  
Direct Submission  
Submitted (03-DEC-1998) Yasukazu Nakamura, Kazusa DNA Research  
Institute, Department of Plant Gene Research; 1532-3, Yana,  
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,  
Tel:81-438-52-3935, Fax:81-438-52-3934)

Address for correspondence: kaos@kazusa.or.jp

For the latest information on annotation of this clone, please see  
http://www.kazusa.or.jp/kaos/cgi-bin/aggd graph.cgi?c=MOB24  
Genes with similarity to proteins in the databases are described in  
'product' or 'note' qualifiers. Genes that have no significant  
protein similarity are described as 'unknown protein'.  
The software programs used to predict genes include: Grail  
(Informatics Group, Oak Ridge National Laboratory,  
http://compbio.ornl.gov/Grail-1.3/),  
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),  
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of  
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and  
SplicePredictor (Volker Brendel, Stanford University,  
http://gremli.zoology.ualb.ca/cgi-bin/sp.cgi).

Genes encoding tRNAs are predicted by tRNAscan-SE  
(Sean Eddy, Washington University School of Medicine, St. Louis,  
http://genome.wustl.edu/eddy/tRNAscan-SE/).

This sequence may not be the entire insert of this clone. It may be  
shorter because we remove overlaps between neighboring submissions.  
The 5' clone is MKP5 and the 3' clone is MSD24.

## COMMENT

Location/Qualifiers

CDS

/note="gene\_id:MOB24.3"  
/codon\_start=1  
/evidence=not experimental  
/product="ethylene-responsive transcriptional  
coactivator-like protein"  
/protein\_id="BAB01997.1"  
/db\_xref="GI:9294040"  
/translation="MPSRYPGANTQDMPEVVLHKSQKSDLRDPKAVNALRNGVAV  
QTVKFDAGNKKGSTAVPVNTKLEETEPEAPMDVRKAEVLMQKARLEKMSQ  
ADAKQNTERTQVVEYENAGQNPQAVLAKMEKVLGVKRGKIGK"  
complement (join(11108..11381,11467..11528,11699..11836,  
11934..12071,12161..12334,12421..12510,12718..13104,  
13623..13765,14306..14405))  
/note="gene\_id:MOB24.4"  
/codon\_start=1  
/evidence=not experimental  
/product="aldhyde dehydrogenase"  
/protein\_id="BAB01998.1"  
/db\_xref="GI:9294041"  
/translation="MENGKNGATTVKLPEIKFTKLFINGOFIDAASGKTETIDPRN  
GEVIATAGGEDVDVLAANAAYAFDHGPMWMTGERAKLNKFDLIEINIELA  
KLDVADGGKLPQGLQYADIATAGHFRYNAGADKHGTETKMTQSLFGYTLKEPIG  
VVGNIIPWNPFSIMFATKVPAMAAGCTVWPKAEQTSLSALFYAHLSEKAGIPDGVL  
NIVTFGSTAGAAIAASHMDVDKVSFTGSTDGRKIMQAAAASNLKVSLELGGKSPLL  
IFNDADIDKAADLALGCFYNKGEICVASSRVFQEGIDKVKVEKAKDVTGDP  
FDSARQGPQVKKQFKILSYIEHKNEGNTLLTGKAKIDGKGYFIQPTIFADVTED  
MKIYDEIFGVSVMKPKTVEGICANNKYGLAAGILSQDIDLINTVSRSLKAGI  
IWNCFYFDLDCPYGGYKMSGNCRESGMDALDNYLQTSVVMPLHNSPMW"  
complement (join(19742..19933,20400..20657))  
/note="gb|AAD25142.1  
gene\_id:MOB24.5  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01999.1"  
/db\_xref="GI:9294042"  
/translation="WTPLSISASPPWPLLSPPLSSSSNLRAAPISFSPSRINHG  
SCSIAVQRALSRTKQLQVAVMAPEEKLTRNPLDPEIWERPKGRPPDIFPK  
FSPMKSFLPPMPYDPPEEEEEKEEETDDPEKEDQPNRS"  
join(23402..23459,23819..24021)  
/note="gene\_id:MOB24.6  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02000.1"  
/db\_xref="GI:9294043"  
/translation="MQNASLKLPLLIFILVITSNLGAARKLTGVADVIVEGEAASPO  
YGIIDENDTLNKSCKRDICTPQCRGFGCNLILOKCELI"  
complement (join(25945..26640,26714..26824,26914..27012,  
27119..27224,27291..27360,27436..27583))  
/note="contains similarity to ubiquitin conjugating enzyme  
gene\_id:MOB24.7"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02001.1"

CDS

/note="gene\_id:MOB24.3"  
/codon\_start=1  
/evidence=not experimental  
/product="ethylene-responsive transcriptional  
coactivator-like protein"  
/protein\_id="BAB01997.1"  
/db\_xref="GI:9294040"  
/translation="MPSRYPGANTQDMPEVVLHKSQKSDLRDPKAVNALRNGVAV  
QTVKFDAGNKKGSTAVPVNTKLEETEPEAPMDVRKAEVLMQKARLEKMSQ  
ADAKQNTERTQVVEYENAGQNPQAVLAKMEKVLGVKRGKIGK"  
complement (join(11108..11381,11467..11528,11699..11836,  
11934..12071,12161..12334,12421..12510,12718..13104,  
13623..13765,14306..14405))  
/note="gene\_id:MOB24.4"  
/codon\_start=1  
/evidence=not experimental  
/product="aldhyde dehydrogenase"  
/protein\_id="BAB01998.1"  
/db\_xref="GI:9294041"  
/translation="MENGKNGATTVKLPEIKFTKLFINGOFIDAASGKTETIDPRN  
GEVIATAGGEDVDVLAANAAYAFDHGPMWMTGERAKLNKFDLIEINIELA  
KLDVADGGKLPQGLQYADIATAGHFRYNAGADKHGTETKMTQSLFGYTLKEPIG  
VVGNIIPWNPFSIMFATKVPAMAAGCTVWPKAEQTSLSALFYAHLSEKAGIPDGVL  
NIVTFGSTAGAAIAASHMDVDKVSFTGSTDGRKIMQAAAASNLKVSLELGGKSPLL  
IFNDADIDKAADLALGCFYNKGEICVASSRVFQEGIDKVKVEKAKDVTGDP  
FDSARQGPQVKKQFKILSYIEHKNEGNTLLTGKAKIDGKGYFIQPTIFADVTED  
MKIYDEIFGVSVMKPKTVEGICANNKYGLAAGILSQDIDLINTVSRSLKAGI  
IWNCFYFDLDCPYGGYKMSGNCRESGMDALDNYLQTSVVMPLHNSPMW"  
complement (join(19742..19933,20400..20657))  
/note="gb|AAD25142.1  
gene\_id:MOB24.5  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01999.1"  
/db\_xref="GI:9294042"  
/translation="WTPLSISASPPWPLLSPPLSSSSNLRAAPISFSPSRINHG  
SCSIAVQRALSRTKQLQVAVMAPEEKLTRNPLDPEIWERPKGRPPDIFPK  
FSPMKSFLPPMPYDPPEEEEEKEEETDDPEKEDQPNRS"  
join(23402..23459,23819..24021)  
/note="gene\_id:MOB24.6  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02000.1"  
/db\_xref="GI:9294043"  
/translation="MQNASLKLPLLIFILVITSNLGAARKLTGVADVIVEGEAASPO  
YGIIDENDTLNKSCKRDICTPQCRGFGCNLILOKCELI"  
complement (join(25945..26640,26714..26824,26914..27012,  
27119..27224,27291..27360,27436..27583))  
/note="contains similarity to ubiquitin conjugating enzyme  
gene\_id:MOB24.7"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02001.1"

## FEATURES

## source

1..79706  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="MOB24"  
/clone\_lib="Mitsui P1"  
2398..3399  
/note="gene\_id:MOB24.1  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01995.1"  
/db\_xref="GI:9294038"  
/translation="MGDSEDTGYPKFYSLNQNHMYSRPIKSHAYVNEEDEDE  
VEGEDPQGGYIRNERFQRQKPNKVVSGFAGPSDAKYADNRQKQAPVLLEW  
GDFRQLGRSLRNDWNEVAEKSEELRMEKSTQCRRMDDLKRYRKIKVKS  
GIGSSKWSFFNNKMLLCCVSPDLGACGVDSGFVFMNTKVYLDKNSGFDEWDS  
GDSDEEDDEVEYERKVNDAASYKMLADSVRFKGYEKMSKKEQMLEKMR  
ADFQDLQKQIIVDRAQSEIARLREEENHHGGDDDEDEMDNSDNLSD"  
complement (join(3794..4276,4363..4490,4577..4673,  
4757..5179,5354..5548,5693..5970,6088..6214,6297..6496,  
6601..6648,6743..6776,6882..7013,7464..7559,7677..7743,  
8162..8217,8290..8385,8470..8998,9122..9462))  
/note="gene\_id:MOB24.2"  
/codon\_start=1  
/evidence=not experimental  
/product="DNA repair protein Muts"  
/protein\_id="BAB01996.1"  
/db\_xref="GI:9294039"  
/translation="MQRKSLISFFQKPTAATTKGLVSGDAASGGGSGGPRFNVKGG  
DAKGDSVRFAVKSVDVVRGTDTPPEKVPVRVLPSPGPKPAESAGDASSLSFNIMHKF  
VKVDRDCSGERSRDEVPLNDSSLCMKANDVLPQPSNNKGTQRNHAFSPSGRAEL

## CDS

/note="gene\_id:MOB24.3"  
/codon\_start=1  
/evidence=not experimental  
/product="ethylene-responsive transcriptional  
coactivator-like protein"  
/protein\_id="BAB01997.1"  
/db\_xref="GI:9294040"  
/translation="MPSRYPGANTQDMPEVVLHKSQKSDLRDPKAVNALRNGVAV  
QTVKFDAGNKKGSTAVPVNTKLEETEPEAPMDVRKAEVLMQKARLEKMSQ  
ADAKQNTERTQVVEYENAGQNPQAVLAKMEKVLGVKRGKIGK"  
complement (join(11108..11381,11467..11528,11699..11836,  
11934..12071,12161..12334,12421..12510,12718..13104,  
13623..13765,14306..14405))  
/note="gene\_id:MOB24.4"  
/codon\_start=1  
/evidence=not experimental  
/product="aldhyde dehydrogenase"  
/protein\_id="BAB01998.1"  
/db\_xref="GI:9294041"  
/translation="MENGKNGATTVKLPEIKFTKLFINGOFIDAASGKTETIDPRN  
GEVIATAGGEDVDVLAANAAYAFDHGPMWMTGERAKLNKFDLIEINIELA  
KLDVADGGKLPQGLQYADIATAGHFRYNAGADKHGTETKMTQSLFGYTLKEPIG  
VVGNIIPWNPFSIMFATKVPAMAAGCTVWPKAEQTSLSALFYAHLSEKAGIPDGVL  
NIVTFGSTAGAAIAASHMDVDKVSFTGSTDGRKIMQAAAASNLKVSLELGGKSPLL  
IFNDADIDKAADLALGCFYNKGEICVASSRVFQEGIDKVKVEKAKDVTGDP  
FDSARQGPQVKKQFKILSYIEHKNEGNTLLTGKAKIDGKGYFIQPTIFADVTED  
MKIYDEIFGVSVMKPKTVEGICANNKYGLAAGILSQDIDLINTVSRSLKAGI  
IWNCFYFDLDCPYGGYKMSGNCRESGMDALDNYLQTSVVMPLHNSPMW"  
complement (join(19742..19933,20400..20657))  
/note="gb|AAD25142.1  
gene\_id:MOB24.5  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01999.1"  
/db\_xref="GI:9294042"  
/translation="WTPLSISASPPWPLLSPPLSSSSNLRAAPISFSPSRINHG  
SCSIAVQRALSRTKQLQVAVMAPEEKLTRNPLDPEIWERPKGRPPDIFPK  
FSPMKSFLPPMPYDPPEEEEEKEEETDDPEKEDQPNRS"  
join(23402..23459,23819..24021)  
/note="gene\_id:MOB24.6  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02000.1"  
/db\_xref="GI:9294043"  
/translation="MQNASLKLPLLIFILVITSNLGAARKLTGVADVIVEGEAASPO  
YGIIDENDTLNKSCKRDICTPQCRGFGCNLILOKCELI"  
complement (join(25945..26640,26714..26824,26914..27012,  
27119..27224,27291..27360,27436..27583))  
/note="contains similarity to ubiquitin conjugating enzyme  
gene\_id:MOB24.7"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02001.1"

## CDS

/note="gene\_id:MOB24.3"  
/codon\_start=1  
/evidence=not experimental  
/product="ethylene-responsive transcriptional  
coactivator-like protein"  
/protein\_id="BAB01997.1"  
/db\_xref="GI:9294040"  
/translation="MPSRYPGANTQDMPEVVLHKSQKSDLRDPKAVNALRNGVAV  
QTVKFDAGNKKGSTAVPVNTKLEETEPEAPMDVRKAEVLMQKARLEKMSQ  
ADAKQNTERTQVVEYENAGQNPQAVLAKMEKVLGVKRGKIGK"  
complement (join(11108..11381,11467..11528,11699..11836,  
11934..12071,12161..12334,12421..12510,12718..13104,  
13623..13765,14306..14405))  
/note="gene\_id:MOB24.4"  
/codon\_start=1  
/evidence=not experimental  
/product="aldhyde dehydrogenase"  
/protein\_id="BAB01998.1"  
/db\_xref="GI:9294041"  
/translation="MENGKNGATTVKLPEIKFTKLFINGOFIDAASGKTETIDPRN  
GEVIATAGGEDVDVLAANAAYAFDHGPMWMTGERAKLNKFDLIEINIELA  
KLDVADGGKLPQGLQYADIATAGHFRYNAGADKHGTETKMTQSLFGYTLKEPIG  
VVGNIIPWNPFSIMFATKVPAMAAGCTVWPKAEQTSLSALFYAHLSEKAGIPDGVL  
NIVTFGSTAGAAIAASHMDVDKVSFTGSTDGRKIMQAAAASNLKVSLELGGKSPLL  
IFNDADIDKAADLALGCFYNKGEICVASSRVFQEGIDKVKVEKAKDVTGDP  
FDSARQGPQVKKQFKILSYIEHKNEGNTLLTGKAKIDGKGYFIQPTIFADVTED  
MKIYDEIFGVSVMKPKTVEGICANNKYGLAAGILSQDIDLINTVSRSLKAGI  
IWNCFYFDLDCPYGGYKMSGNCRESGMDALDNYLQTSVVMPLHNSPMW"  
complement (join(19742..19933,20400..20657))  
/note="gb|AAD25142.1  
gene\_id:MOB24.5  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01999.1"  
/db\_xref="GI:9294042"  
/translation="WTPLSISASPPWPLLSPPLSSSSNLRAAPISFSPSRINHG  
SCSIAVQRALSRTKQLQVAVMAPEEKLTRNPLDPEIWERPKGRPPDIFPK  
FSPMKSFLPPMPYDPPEEEEEKEEETDDPEKEDQPNRS"  
join(23402..23459,23819..24021)  
/note="gene\_id:MOB24.6  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02000.1"  
/db\_xref="GI:9294043"  
/translation="MQNASLKLPLLIFILVITSNLGAARKLTGVADVIVEGEAASPO  
YGIIDENDTLNKSCKRDICTPQCRGFGCNLILOKCELI"  
complement (join(25945..26640,26714..26824,26914..27012,  
27119..27224,27291..27360,27436..27583))  
/note="contains similarity to ubiquitin conjugating enzyme  
gene\_id:MOB24.7"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02001.1"

/db\_xref="GI:9294044"  
/translation="MAQAARLSLRMOKELKLLSDPPHAGSFPHLSAANGSGDFSTF  
STIDAOLEGPDVTYANGIENLKIQIPERYFOPPIIVSFATPIVHPNIDNSGRICLDI  
LANLPKGNQWQSLNISTVLTSRMLLLSEPNDDGLMCEVSREYKXNQTFDYKAREMT  
EDNGKRAVVGFSGEGNSFGNDGKRSKSLSLALPSQKQKJCGEELTRGVSAAC  
KENKPNLNGKLSLGLKQPCNDTLAGPRTSAKSDNNRLSRKLSLSPGLGELNEVSK  
PEVLAQTDMLKMNQNDARSURGFENFVLEETSMASIVVLSDSDSGQBEERBSVS  
SRSLSLAKRRVLAKRP"

CDS

28636..30168  
/note="emb|CAB79678.1  
gene\_id:MOB24.8  
similar to unknown protein"  
/evidence=not experimental  
/protein\_id="EAB02002.1"  
/db\_xref="GI:9294045"

/translation="MSSLIAYLSFNSFPLKLSRILSNLWLVRTSTNKKRWCL  
AIPTYHLHATPLWKSSITMLLRQFSWVEPTTYRFEISLRAFILTQAMSHGETLLWD  
TTSFLGLTRILWFSSASPHQONQSTTFSPENDASLSSPYNSRHHDLHLSEGLK  
GLQVLGLVLPGLRYMPI SKDDPKLPKPSDLNYSLSIFQSLSVHLASIDRHH  
TNHNLVLIWLHKNSLRTPLVSSRTFWTCGICLKNGIMPTTTSRWCDLCHPSATI  
CISTHLGISFKQSCISRVKAHSRISKEVPNLFETIPEPSHALVVGKLGKIGMI  
TSPSGDYQRFNLSHSPQITTEPKQRFNLIKFSIHDSLGKNGIMISPMWSDVVR  
SFFNLDLPTIVTHLGDQQLLVARSARTSALEPYSTSRLLITVNLASIDSFVED  
PSTTRDLTCAIMPYSLSLWKLALNMFMSFHIIVHLVLGNASPNYVLFVIFESLLLIIV  
LI"

CDS

join(32831..33058,33143..33907)  
/note="contains similarity to heat shock transcription  
factor  
gene\_id:MOB24.9"

/codon\_start=1  
/evidence=not experimental  
/protein\_id="EAB02003.1"  
/db\_xref="GI:9294046"

/translation="MEDDNNNNNNNNVIAPIVTKYQMVNDPSTDWLIITWGPANHNSF  
IVDPLDSQRIILPILPKHNNFSSVRLQNTYFRKVDPRWEFANFHLRGQKHLIN  
NIARKKARGMYGQDEGEIVREIRLKEORELEAEIORMNRRIEATEKRPMQMA  
FLYKVVDPDLLPMLLEKERTKQOQVSDKKERTVMSTVKSEEVERDEGRVFRV  
MSSSTPSPSTENLYRHSPPGWIIVPMTQQPGSYETGLVAKSLNSSTSTSSSLTS  
TFSUPESVGGGGGGGSIQGERIKETATFGVVENPPPTTPPYPSLSFRGGH"  
complement(join(37054..37590,37670..37809,37894..38197,  
38288..38398,38486..38584,38662..38772,39019..39066,  
39311..39367))  
/note="gene\_id:MOB24.10  
unknown protein"  
/codon\_start=1

CDS

## Alignment Scores:

Pred. No.:	7,29e-55	Length:	79706
Score:	2622.00	Matches:	573
Percent Similarity:	61.19%	Conservative:	23
Best Local Similarity:	58.83%	Mismatches:	43
Query Match:	75.93%	Indels:	337
DB:	8	Gaps:	19

US-10-086-464-2 (1-647) x AB020746 (1-79706)

Qy	1	MetSerSerAlaProSerProGlyThr-----GlySerProProSerProSer 17
Db	51787	ATGTCACAGCGCGTCTCCAGGCACTACTCCATCACCCTCCACCGCTCTCTCCACCA 51846
Qy	18	AnSerThrThrThrProProAla-----SerAlaProProThrThrPro 35
Db	51847	AACTCGACACACCACTCTCTCCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 51906
Qy	36	SerSerProProPro---ProSerThrThrThrThrThrThrThrThrThrThrThr 54
Db	51907	TCCTTCT 51966
Qy	55	ThrProSerAlaProProProSerProProThrThrThrThrThrThrThrThrThr 74
Db	51967	CCTCCTTCT 52014

Qy	75	LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94
Db	52015	CTTCTCAACTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 52071
Qy	95	ProProThr---ArgAenProProProSerValProGlyProProSerAenProSerArg 113
Db	52072	CCCTCAAAACCTCGAAGCCCTCCATCTCTAAACCAAGGACCAACCAACACTCCCTCA--- 52128
Qy	114	GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131
Db	52129	---GGATCTACTCTAGAACCTCCATCAACACTAAACCGTCGCGCGCTCTGATCTCTCC 52185
Qy	132	AspGlyLeuSerThrGlyValValValGlyLeuAlaIleGlyGlyValAlaLeuVal 151
Db	52186	GATGATGTTCTACCGGAGTTGTGTAGGAATGCCATTTGGAGGATCGCTATCTTCTGTT 52245
Qy	152	IleValThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169
Db	52246	ATACTGACTCTGATTTGCTTCTCTCTGTAAGAAGAAACGAAGAAGACACACGACGATA 52305
Qy	170	AspAlaTyThrValProProProProProProProProProProProProProProPro 184
Db	52306	GCTGCTTACTATGTTCT 52364
Qy	184	----- 184
Db	52365	CAACTTCATGAATAAATAATCGAATATTTAACCATAGTTGTTCAATTTCTGATCGAATTG 52424
Qy	184	----- 184
Db	52425	CTCAATCTGATCTTCTGATTATGCTTATCTGTTTAAAGAGATTACCTTTGTTGGAATA 52484
Qy	185	----- 185
Db	52485	GTTTCAAACTTTACAGAGATTGATGTTGTTTATTCAGCTCGAGGACCTTACGGTGGTCAA 52544
Qy	192	GlnGlnGlnTrpArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 210
Db	52545	CAACAGTATTGGACAAACAAACGCGTCACGGCGTCAGATATCATGATGATGATGATGAT 52604
Qy	211	LeuProProProProProProProProProProProProProProProProProProPro 230
Db	52605	TTG---CCACCACCTAAGCTCTCCATCTCCACCACGAAACCTCTCTCCGCCACCTCCACA 52661
Qy	231	Pro---PheMetSerSerSerGlyGlySerAspTyThrSerAspArgProValLeuProPro 249
Db	52662	CCAGCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 52721
Qy	250	ProSerProGlyLeuValLeuLeuGlyPheSerLysSerThrPheThrTyThrGluLeuAla 269
Db	52722	CCATCTCCAGGCTTGTTGTTAGGCTTTTCTAAAAGCACTTTCACTTAAGAGGATGTTGTCG 52781
Qy	270	ArgAlaThrAenGlyPheSerGluAlaAenLeuLeuGlyGlnGlyGlyPheGlyTyxVal 289
Db	52782	AGAGTACTAATGGCTTCTCTGAGGCTAATTTGTTAGGACAAAGGAGGTTGTTGTTATGTG 52841
Qy	290	HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
Db	52842	CATAAAGTATATTGCTCTAGTGGGAAAGATTGCTGTGAAACAGATTGAAAGCTGGTAGT 52901
Qy	310	GlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuIleSerArgValHisHisArg 329
Db	52902	GGTCAGGAGAGAGAGATTTCAGGCTTGAGGTTGAGATCATTAGCAGAGTTTCATCAGG 52961
Qy	330	HisLeuValSerLeuValGlyTyThrCysIleAlaGlyAlaLysArgLeuLeuValTyxGlu 349
Db	52962	CAITTGCTTCTCTTATGTTGTTATGATGCGCGGTGTTCAAGATTACTTGTCTATGAG 53021
Qy	350	PheValProAenAenAenLeuLeuLeuHisLeuHisGlyGlu----- 363
Db	53022	TTTGTTCCAAAACAACATCTTGAGTTTCACTCCATCGTTTA-GAACAACAACTAACTCTT 53080

QY 363 ----- 363  
Db 53081 TCCTTTGATTTGTTACATGCTTGTGTTGTTTTCAGTTAGTGAAGTTGTTGGGAATT 53140  
QY 364 -----GlyArgProThrMetGluTrpSerThrArgLeuLysIleAla 377  
Db 53141 GNAATTTGATAGGTAAGGACGGCTACGATGGATGGATGAGTACTAGATTGAAGATTGCT 53200  
QY 378 LeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys----- 391  
Db 53201 CTGTGATCTGCTAAAGGACTTTCATATCTTCATGAAGATTCTAAGTCTAAACCTTAAAA 53260  
QY 391 ----- 391  
Db 53261 ATTATGTACAGATGTTTGATAAATCTCTCTTGTGTTCTCACTGAGGTATATATATAAAT 53320  
QY 392 -----AenProLysIleIleHisArgAspIleLysAlaSerAenIleLeuIleAs 408  
Db 53321 ATTTCGAGGCAATCCGAANAATCATTCACCGTGATATTAAAGGCGTCAACATATTGATTGA 53380  
QY 408 pPheLysPheGluAlaLys----- 414  
Db 53381 TTTCAAATTTGAAGCTAAGGTATGTTTACTCGATATGTTTGTGAAGTGTGTTTGTGAATT 53440  
QY 415 -----ValAlaAspPheGlyLeuAlaLysIleAlaAs 425  
Db 53441 TGTGTCACTAAAGCTCTTCTTCTTTCAGGTTGCTGACTTGTGCTTGTCCGAAGATTGCTT 53500  
QY 425 erAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly----- 440  
Db 53501 CTGATACAAACNCTCATGTAICTACACCGGTGATGGGAACCTTTGGGTAAAAATCATCTT 53560  
QY 440 ----- 440  
Db 53561 TAACTACTAAACATGTACATTTATGTCAGTTGCTTAAATGTCATGTGATCTTGTGCAACA 53620  
QY 441 -----TyrLeuAlaProGluTyrAlaAlaSerG 450  
Db 53621 AGTTCGTGAACCTTTTGGTTTGGGATTTTAGGTATTTGGCTCCGGAATATGCTGCAAGTG 53680  
QY 450 lYlYsLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleT 470  
Db 53681 GAAAGCTCACAGAAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGGAACTTATTA 53740  
QY 470 hrGlyArgProValAspAlaAenValTyrValAspAspSerLeuValAspTrp- 489  
Db 53741 CTGGAGGCGCCCTGTTGATGCGAAACAATGCTATGTAGATGACACAGCTTAGTTGACTGGG 53800  
QY 489 ----- 489  
Db 53801 TACTTTTACTATGCTGCAGATAAATCTAAACTTCTGATTTTCTACAATTTAACTGTTGAC 53860  
QY 490 -----AlaArgProLeuLeuAenArgAla 497  
Db 53861 TAAACCAAAAGCTATGTGTGATTGATGTTGGTGGNACAGGACGACCATTCCTTAAACGACGA 53920  
QY 498 SerGluGlnGlyAspPheGlyLeuAlaAspAlaLysMetAsnAenGlyTyrAspArg 517  
Db 53921 TCTGAGGAAGGAGATTTTGGGGTTTGGCTGATTCAAAGATGGGTAATGAGTATGACAGA 53980  
QY 518 GluGluMetAlaArgMetValAlaCysAlaAlaAlaCysValAlaTqHisSerAlaArgArg 537  
Db 53981 GAGGAGATGGCTCGCATGTTGCTTGGCGTGGCGTGTGTTGTCGCAATTCAGCTCGCCGC 54040  
QY 538 ArgProArgMetSer----- 542  
Db 54041 AGACCTCGCATGAGCGCAGGTAACCTTAATCTTCTCTAGAGAAGCATATTGCTTTGG 54100  
QY 542 ----- 542  
Db 54101 TTCTGTTTATTTGGTAGAATAATATGTTTCCGTATGTTTAAAGAAAGGTTTGTATGATG 54160  
QY 542 ----- 542

Db 54161 GTTTTTGTGAGAAAGATCGGTTCAGTTATGTTCTTATGCTAGCCCTACTTTTTGTTTACC 54220  
QY 542 ----- 542  
Db 54221 CATCATCATACTATCTAAACAACACCGATGGAATATGTCAAAGAATTTGGCTGCAGTTTC 54280  
QY 542 ----- 542  
Db 54281 TTTTGTGTTTTCATTTTGGCAAGGATTGTCAATTTTAAACATATGTTAATGCTAATGGAGTGAG 54340  
QY 543 -----GlnIleValArgAlaLeuGluGlyAsnValSerLeuSer 555  
Db 54341 TGATTTGCTGTACAAAATAGCAGATAGTAGCGGGGTTAGAGAAGAAATGATCGCTGTCT 54400  
QY 556 AspLeuAenGluGlyMetArgProGlyGlnSerAenValTyrSerSerTyrGlyGlySer 575  
Db 54401 GATCTTAAACGAAGGATGAGACCGGTCACAGCAACGTATACAGCTCATATGGAGGAAGC 54460  
QY 576 ThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeu 595  
Db 54461 ACAGACTATGACACGAGCCAAATACACGACGACATGATAAAGTTTAGGAAATGGCTCTT 54520  
QY 596 GlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAenProThrSerAspTyrGlyLeu 615  
Db 54521 GGAATCTAAGNATACGACACACCGGCGAGTACAGTAATCCACACAGTACTCGGACTG 54580  
QY 616 TyrProSerGlySerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIle 635  
Db 54581 TACCGGCTGTTGGTTCAAGCAGTGAAGGTCAAGCCACACGAGAAATGGAGATGGAAAGATT 54640  
QY 636 LysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647  
Db 54641 AAGAAACCCGTCGAAGTTATAGTGGACCCCTCTCTT 54676  
RESULT 15  
AKI03247  
LOCUS  
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033123K23, full insert sequence.  
ACCESSION AKI03247  
VERSION AKI03247.1 GI:32988456  
KEYWORDS FLI CNA; CAP trapper.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1  
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team.; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shihahiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otono,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kuusmaki,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN.; Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Oka,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.  
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
JOURNAL Science 301 (5631), 376-379 (2003)  
MEDLINE 22752273  
PUBMED 12869764  
REFERENCE 2 (bases 1 to 2307)  
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,



Job time : 5986 secs

```
Db 988 TATGTTCAAGAGGAGTTCTGCCGAATGGACAGAGGTGCTGTGAACAAATTGAGAGAT 1047
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArgValHis 327
Db 1048 GGGAGTGGCAGGAGAGCGTGAATTTCAGCGGAGGTTGAGATTATCAGCGGGTACAT 1107
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347
Db 1108 CACAAGCATCTGTAAACATGGTTGGTTATTGTCATTTCTGGAGGAGAGAGGTTGCTTCTC 1167
Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThr 367
Db 1168 TATGAGTATGTTCCAAAACAACATATGGAGCTCCACTTGGCATGGGAGAGCGGACCAACC 1227
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387
Db 1228 ATGGATGGCCTACAAAGATTACGCATTGCTTTGGGTGCTGCGAAGGATTGGCATACTT 1287
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407
Db 1288 CATGAAGATTGCCATCTCAAGATCATTCACCGTGACATAAAGTCAGCAAAATATTCTTCTT 1347
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427
Db 1348 GACCGGAGATTGGGCAAAAGTGCGAGATTTCGGACTTCGAAATTCGACTTCGTGATAAT 1407
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
Db 1408 AACACTCATGTTTCAACACAGAGTAATGGGCACATTTGGGTACCTTGCACAGAGATGCG 1467
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db 1468 TCTTCTGGCCAGCTAACTGAGAAATCAGATGTCCTTTCTTCGGAGTAATGCTTCTTGAG 1527
Qy 468 LeuIleThrGlyArgProValAspAlaAsnAsnValTyrValAspAspSerLeuVal 487
Db 1528 CTAATAACTGGCCGCCGCTCTGTGAGATCGAAACCAATCACAGATGGATGACAGCTTGGTT 1587
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db 1588 GATTGGGCAAGGCTTTAATGATGCGAGCATCAGATGATGGCAATTATGATGCTTTAGTG 1647
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db 1648 GATCCACGGCTAGGACAGAGTACAAATGGTAACGAGATGGCGAGATGATTGCTTGTGCA 1707
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db 1708 GCTGCTTGTGTACGCCATTCTGCACGCCGACGCCACCGAGTGGCGGATGTTGTCGGGCC 1767
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsn 567
Db 1768 TTGGAAGCGGATGTCATCGATGATCTGAACGAAGGTGTTCTGCTCGCCATAGCCGA 1827
Qy 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587
Db 1828 TTCTTGGGATCATAC---AACAGCAACGAGTACGACACGGGCCATTACAATGAAGACCTG 1884
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSer 607
Db 1895 AAGAAATTCCAGGAAGATGGCATTGGCAGTGGCAATCTA-----GAGAGCAGC 1932
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
Db 1933 CAGCAACGCGAGCCCGAGTTCGTCGGAATCGATCGTATCAATGGGGGATGCGCGT 1992
Qy 628 Arg-----GluMetGluMetGlyLysIleLysArgThrGly 639
Db 1993 CAAATACCGGACGAGATGGAGTGGGAGTTTGAAGAAAGATGCG 2040
```

Search completed: May 13, 2004, 00:33:01

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2004, 20:46:26 ; Search time 603 Seconds  
(without alignments)  
4558.183 Million cell updates/sec

Title: US-10-086-464-2

Perfect score: 3453

Sequence: 1 MSSAPSGTGTGPPSPPSNST.....REMEMGKIKRTGQYSGPSL 647

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp  
-Q=/cgn2.1/USPTO\_spool\_p/US10086464/runat\_11052004\_121814\_7833/app\_query.fasta\_1.839  
-DB=N Geneseq\_29Jan04 -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORF=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10086464 -CGN\_1\_1\_708 -runat\_11052004\_121814\_7833 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	1944	4	Aaf77094 Brassica
2	3453	100.0	1944	9	AcF36548 B. napus
3	3453	100.0	2189	9	AcF36549 B. napus
4	2936.5	85.0	1959	9	AcF36557 A. thaliana
5	2885	83.6	2324	3	AcF36568 Arabidops
6	1821.5	52.8	1812	9	AcF36554 O. sativa
7	1668	48.3	1515	9	AcF36555 Z. mays p
8	1637.5	47.4	1488	9	AcF36556 G. max PE

9	1507	43.6	2520	4	Aaf77096	Arabidops
10	1504	43.6	2520	9	AcF36551	Arabidops
11	1479.5	42.8	2820	4	Aaf77097	Arabidops
12	1471.5	42.6	2820	9	AcF36552	Arabidops
13	1209.5	35.0	2880	4	Aaf77095	Arabidops
14	1209.5	35.0	2880	9	AcF36550	Arabidops
15	1185	34.3	3060	9	AcF36553	Arabidops
16	1181	34.2	3060	4	Aaf77098	Arabidops
17	839.5	24.3	1923	7	ADA70890	Rice gene
18	816.5	23.6	2124	6	ABZ13274	Arabidops
19	804.5	23.3	2865	7	ADA70701	Rice gene
20	802.5	23.2	2011	4	AAH22557	P. patens
21	797	23.1	1730	7	ADA70647	Rice gene
22	793.5	23.0	2823	7	ADA69478	Rice gene
23	782.5	22.7	1419	3	AAC43029	Arabidops
24	782.5	22.7	1419	6	ABZ13121	Arabidops
25	775.5	22.5	1626	3	AAC48313	Arabidops
26	775.5	22.5	1917	7	ADA70527	Rice gene
27	773	22.4	1479	3	AAC43040	Arabidops
28	769.5	22.3	1353	3	AAC42860	Arabidops
29	769.5	22.3	1353	6	ABZ14106	Arabidops
30	766	22.2	1494	3	AAC48560	Arabidops
31	758.5	22.0	1572	3	AAC42898	Arabidops
32	758.5	22.0	1572	6	ABZ13705	Arabidops
33	758.5	22.0	1920	6	AAL45905	A thalian
34	758.5	22.0	2012	3	AAC44847	Arabidops
35	756	21.9	3239	3	AAZ79271	Eucalyptu
36	754.5	21.9	2136	7	ADA69829	Rice gene
37	749.5	21.7	1751	3	AAC45313	Arabidops
38	744.5	21.6	1485	3	AAC42702	Arabidops
39	740.5	21.4	3525	7	ADA71170	Rice gene
40	740	21.4	2868	3	AAZ79266	Pinus rad
41	740	21.4	3153	7	ADA71066	Rice gene
42	736.5	21.3	1703	3	AAC46629	Zea mays
43	736	21.3	3624	7	ADA69512	Rice gene
44	735	21.3	2004	3	AAC40780	Arabidops
45	734	21.3	1455	3	AAC45877	Arabidops

#### ALIGNMENTS

RESULT 1

AAf77094  
ID AAF77094 standard; DNA; 1944 BP.

XX AC AAF77094;

XX DT 17-MAY-2001 (first entry)

XX DE Brassica napus PERK1 DNA.

XX KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.

XX OS Brassica napus.

XX PN WO200114563-A1.

XX PD 01-MAR-2001.

XX PF 18-AUG-2000; 2000WO-CA000966.

XX PR 19-AUG-1999; 99US-0149466P.

XX PR 13-OCT-1999; 99US-0159122P.

XX XX (GORI/) GORING D.

XX PA (SILV/) SILVA N.

XX PI Goring D, Silva N;

XX XX WPI; 2001-244305/25.

XX XX New proline-rich, extensin-like receptor kinase nucleic acids and  
PT polypeptides useful for increasing plant wounding or pathogen resistance,

PT or for producing transgenic plants with increased wounding or pathogen  
 PT resistance.

PS Claim 6; Fig 1; 91pp; English.

XX The present invention relates to proline-rich extensin-like receptor  
 CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for  
 CC increasing the resistance of plants to wounding and pathogens. These are  
 CC also useful for producing transgenic plants with increased wounding and  
 CC pathogen resistance compared with a wild type plant, as well as in assays  
 CC for identifying and developing compounds to inhibit and/or enhance  
 CC polypeptide function directly

XX Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 4, 62e-93 Length: 1944  
 Score: 3453.00 Matches: 647  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-086-464-2 (1-647) x AAF77094 (1-1944)

QY 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20  
 DB 1 ATGCTCTCGGCGCGCTCTCCGGGACTGGTTCGCCCTCCATCTCCACCATCAAACTCCACA 60  
 QY 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProPro 40  
 DB 61 ACCACCATCTCTCTCCAGCTTCGGCTCCTCTCCACACACACCTCTCTCTCTCCGCGG 120  
 QY 41 ProSerThrIleProThrSerProProSerSerArgSerThrProSerAlaProPro 60  
 DB 121 CCATCCACTATTCGAGATCT 180  
 QY 61 ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80  
 DB 181 CCATCTCCACCAACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCCA 240  
 QY 81 ProAlaProThrThrProGlySerProAlaProValThrProProThrArgAsnPro 100  
 DB 241 CCCGCTTCAACTACCGCCGGATCTCCACCGCACTGTCTCTCTCTCTCTCTCTCTCTCTCT 300  
 QY 101 ProProSerValProGlyProProSerAsnProSerArgGlyGlySerProArgPro 120  
 DB 301 CCACCTTCAGTCCAGGACCCCGTCCATCTCTTCCAGGAGGATCTCTCGACCT 360  
 QY 121 ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValVal 140  
 DB 361 CCATCTTCTCCCTCGCGCGCT 420  
 QY 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysIleLeuCys 160  
 DB 421 GGAATCGGCATCGGAGGAGTCTGCTCTGTGTGAGTAGTACTCTGATTGTCTCTCTCTCT 480  
 QY 161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProPro 180  
 DB 481 AAGAAGAAGACGCGAG 540  
 QY 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnTrrArgGlnGlnAsnAla 200  
 DB 541 GGTCTCCAAAGCGGAGGACCTTACGGTGGACAGCAGCAACAATGCGGCGCAACAAACGCA 600  
 QY 201 ThrProProSerAspHisValValThrSerLeuProProProProProProProPro 220  
 DB 601 ACACCACCGTCAGATCATGTGTGACGTCACTTACCACCAACCACTTAAGGCTCCATCTCCA 660  
 QY 221 ProArgGlnProProProProProProProPheMetSerSerSerGlyGlySerAsp 240  
 DB 661 CCACGGCAACCTCTCTCCACCTCCACCGGCTTTCATGAGCAGCAGCGGCGGCTCCGAC 720

QY 241 TyrSerAspArgProValLeuProProSerProGlyLeuValLeuGlyPheSerIys 260  
 DB 721 TACTCGGACCGTCCAGTTCTTCTCCACCGCTCTCCAGGCGTGTGTAGGCTTCTCCAAA 780  
 QY 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280  
 DB 781 AGCACTTTCACATACAGGAGGCTAGCTAGAGCCCACTAATGGTTCTCCGAGCGCAACTTG 840  
 QY 281 LeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProSerGlyGlyGluVal 300  
 DB 841 TTAGGACACAGCGGGTTCGGTTACGTGCACAAAGTGTGTTCCTAGTGGGAAAGAGATT 900  
 QY 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320  
 DB 901 GCTGTGAAGCAGTTGAAGTTGGAGTGGTTCAGGAGAGAGGAGTTCAGCGAGAGGTT 960  
 QY 321 GluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340  
 DB 961 GAGATCATCAGCAGATTCCACACAGCACTGGTGTCTCTTGTGGTATTATGATCGCC 1020  
 QY 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360  
 DB 1021 GGTGCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAACAACAATCTCGAGCTTCACCTC 1080  
 QY 361 HisGlyGluGlyArgProThrMetGluThrSerThrArgLeuLysIleAlaLeuGlySer 380  
 DB 1081 CATCGCAGGACCGCTTACAATGGAATGGAGCCAGATTGAAGATTGCTTCTTGGATCT 1140  
 QY 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400  
 DB 1141 GCTAAAGCATCTTCTTATCTTCATGAAGTTGCAATCTTAATAATCATTCACCGTGATATC 1200  
 QY 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420  
 DB 1201 AAGGCTTCAACAATATTGATAGATTTCAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT 1260  
 QY 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440  
 DB 1261 GCTAAGATTGCTTCTGATACAAACACCATGTATCAACACGTGTGTGGGAACTTTGGG 1320  
 QY 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460  
 DB 1321 TACTTGGCTCGGAATACGCTGCAAGCGGAAGCTCACGGAGAGTCTGACGTTTCTCA 1380  
 QY 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480  
 DB 1381 TTTGGCGTTGTCTTTTGGAGCTCATTTACTTGAGCTCGACCGTCCGACCAATGTC 1440  
 QY 481 TyrValAspAspSerLeuValAspTrrAlaArgProLeuLeuAsnArgAlaSerGluGln 500  
 DB 1441 TATGTAGATGACAGCTTAGTTGACTGGGCGACGACCATTTGCTTAACCGAGCATCTGAGCA 1500  
 QY 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520  
 DB 1501 GGAGACTTTGAGGGTTAGCTGATGCAAGAGATGAATATGCGTATGACAGAGAGAGATG 1560  
 QY 521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540  
 DB 1561 GCTCGCATGGTGTGCTGTGCTCGGCTTGTTCGCCATTTCAGCTCGCGCGAGACCTCGC 1620  
 QY 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560  
 DB 1621 ATGAGCCAGATTGTGCGTGTGAGGAAATGTATCCTGTGTGATGCTTTAACAAGAGG 1680  
 QY 561 MetArgProGlyGlnSerAsnValTyrSerTyrGlyGlySerThrAspTyrAspSer 580  
 DB 1681 ATGAGCAGGCTCAAGCAATGTATACAGCTCATACGGAGGAGACCGATTAATGACTCG 1740  
 QY 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600  
 DB 1741 AGCCAGTACAAATGAAGACATGAGAAGTTTAGGAAATGGCACTTGGAACTCAAGAGTAC 1800  
 QY 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620



Db 901 GCTGTGAAGACGTTCAAAAGTTGGAGTGGTCAGGAGAGAGGAGTTTCAGGCAGAGGTT 960  
Qy 321 GluLeIleSerArgValHiHisArgHisIleuValSerLeuValGlyTyrCysIleAla 340  
Db 961 GAGATCATCAGCAGAGGTTCCACACAGGCATCTGGTGTCTCTTGTGGTTATTGTCATGCC 1020  
Qy 341 GlyAlaIysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeu 360  
Db 1021 GGTGCCAAAAGATTGCTTGTCTATGAGTTTGTCTTAAACAATCTCGAGCTTCACCTC 1080  
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380  
Db 1081 CATGGCGAGGACCGCCTACATGAATGGAGCACCAGATTGAAGATTGCTCTGGATCT 1140  
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400  
Db 1141 GCTAAAGGACTTCTTATCTTCATGAAGATTGCAATCTTAAATCATTCACCGTGATATC 1200  
Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420  
Db 1201 AAGGCTTCAACATATTATGATTTCAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT 1260  
Qy 421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440  
Db 1261 GCTAAGATTGCTTCTGATACAAACACGCGATGATCAACACGCTGTGATCGGAACCTTTGG 1320  
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460  
Db 1321 TACTTGGCTCGGAATAGCTGCAAGCGGAAGCTCACGAGAGTCTGAGCTTTTCTCA 1380  
Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnVal 480  
Db 1381 TTTGGCGTGTGCTTTTGAGCTCATTTACTGACGTGACCCGTTGATGCCCAACAATGTC 1440  
Qy 481 TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500  
Db 1441 TATGTAGATGACAGTTAGTTGACTGGCAGCAGCAATGCTTTAACCGAGCATCTGAGCAA 1500  
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluMet 520  
Db 1501 GGAGACTTGGGGTTAGCTGATGCCAAAGATGAATAATGGGTATGCACAGAGAGAGATG 1560  
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaAtcArgArgProArg 540  
Db 1561 GCTCCATGGTGTGCTGTGCTGCGGCTTGTGTTCGCCATTGAGTTCGCCGAGACCTCGC 1620  
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560  
Db 1621 ATGAGCCAGATTGTCGTGCGTTAGAGGAATGTATCACTGTCAAGATCTTAACGAAGG 1680  
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580  
Db 1681 ATGAGACCAGGTCAAAAGCAATGTATACAGCTCATACGGAGGAAGCACCATTATGACTCG 1740  
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600  
Db 1741 AGCCAGTACAATGAAGACATGAAGAAAGTTTAGGAAAATGGCACCTGGAACTCAGAGATC 1800  
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620  
Db 1801 AACGCCACGGGTGAGTACAGTAATCCGACCATGACTATGAGCTGTACCCGCTCGTTCA 1860  
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640  
Db 1861 AGCAGCAGGGGCCAAACCAACACACGCAATGGAGATGGGGAAGATTAAAGAAACCGGTCA 1920  
Qy 641 GlyTyrSerGlyProSerLeu 647  
Db 1921 GGTATAGTGGACCTTCTCTT 1941

RESULT 3  
ACF36549  
ID ACF36549 standard; cDNA; 2189 BP.

XX ACF36549;  
AC 18-DEC-2003 (first entry)  
DE B. napus PERK1 receptor kinase encoding cDNA.  
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;  
KW wound; pathogen resistance; plant growth; seed production; gene; ss.  
XX Brassica napus.  
XX Key Location/Qualifiers  
CDS 1..2189  
FT /\*tag= C  
FT /note= "ABR82938; this protein contains the amino acids  
FT corresponding to 5' and 3'UTR regions, though only the  
FT relevant aa residues (ABR82937) is used in the invention"  
FT 5'UTR 1..96  
FT /\*tag= a  
FT CDS 97..2040  
FT /\*tag= b  
FT /transl\_except= (pos: 1332..1334, aa: Phe)  
FT /product= "PERK1"  
FT /note= "ABR82937; proline-rich Extensin-like Receptor  
FT Kinase"  
FT 3'UTR 2041..2189  
FT /\*tag= d  
XX WO2003072763-A1.  
XX 04-SEP-2003.  
XX 28-FEB-2003; 2003WO-CA000274.  
XX 28-FEB-2002; 2002CA-02373903.  
XX 28-FEB-2002; 2002US-00086464.  
XX (GORI/) GORING D.  
XX (SILV/) SILVA N.  
XX (HAFF/) HAFFANI Y Z.  
XX Goring D, Silva N, Haffani YZ;  
XX WPI; 2003-712727/67.  
XX P-ESDB; ABR82937, ABR82938.  
XX Producing a transgenic plant having an increased plant resistance, plant  
XX growth or seed production comprises transforming a plant with a nucleic  
XX acid molecule having a Proline-rich Extensin-like Receptor Kinase  
XX activity.  
XX Disclosure; Fig 1D; 123pp; English.  
XX The invention relates to producing a transgenic plant having increased  
XX plant height, number of branches, number of seed pods and/or seed  
XX production compared to a non-transgenic plant, and/or quicker flowering  
XX or later senescence compared to a non-transgenic plant. The method  
XX involves transforming a plant with a vector including a proline-rich  
XX Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic  
XX acid molecule having PERK activity. The method, as well as the PERK  
XX nucleic acid molecule and polypeptide, are useful in increasing plant  
XX resistance to wounding and pathogens, and in increasing plant growth and  
XX seed production. The nucleic acid molecule and polypeptide may also be  
XX used in producing transgenic plants or transgenic host cells. The present  
XX sequence represents a cDNA encoding a B. napus PERK1 receptor kinase  
XX polypeptide  
XX Sequence 2189 BP; 544 A; 592 C; 494 G; 559 T; 0 U; 0 Other;  
SQ

Alignment Scores:  
Pred. No.: 5e-93 Length: 2189  
Score: 3453.00 Matches: 647

Percent Similarity: 100.00%    Conservative: 0  
Best Local Similarity: 100.00%    Mismatches: 0  
Query Match: 100.00%    Indels: 0  
DB: 9    Gaps: 0

US-10-086-464-2 (1-647) x ACF36549 (1-2189)

Qy	1	MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr	20
Db	97	ATGTCCTCGGCGCGCTCTCGGGGACTGGTTGCGCTCTCCATCTCCACCATCAAACTCCACA	156
Qy	21	ThrThrThrProProAlaSerAlaProProProProThrThrProSerSerProProPro	40
Db	157	ACCACCACTCCTCTCCAGTCTCGCTCTCCCTCCACCACTCTCTCTCTCTCTCTCTCTCT	216
Qy	41	ProSerThrThrProProSerProProSerSerArgSerThrProSerAlaProPro	60
Db	217	CCATCCACTATTCGACATCT	276
Qy	61	ProSerProProThrProSerThrProGlySerProProProLeuProGlnProSerPro	80
Db	277	CCATCTCCACCACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCTCTCT	336
Qy	81	ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro	100
Db	337	CCCGCTCCAACTAGCGCGGATCTCCACCGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	396
Qy	101	ProProSerValProGlyProProSerAsnProSerArgGlyGlySerProArgPro	120
Db	397	CCACCTTCAGTCCCGGACCACT	456
Qy	121	ProSerSerProProProSerProSerAspGlyLeuSerThrGlyValValVal	140
Db	457	CCATCT	516
Qy	141	GlyThrAlaThrGlyValAlaLeuLeuValThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	160
Db	517	GGAAATCGCATCGGAGAGTCT	576
Qy	161	LysLysLysArgArgArgAspGluAspAlaTyrValProProProProProProProPro	180
Db	577	AAGAAGAACACCGAGAGACGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	636
Qy	181	GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	200
Db	637	GGTCCCAAGCGGAGGACCTTACGGTGGACAGAGCAACATGGCGGCAACAAACGCA	696
Qy	201	ThrProProSerAspHisValValThrSerLeuProProProProProProProProProPro	220
Db	697	ACACCACCGTCAGATCATGTCGTGAGGTCACTACCCACCACCTAAGGCTCCATCTCCA	756
Qy	221	ProArgGlnProProProProProProProProProProProProProProProProProPro	240
Db	757	CCACGGCAACCTCTCCACCTCCACCGCTTTTCATGACGACGCGCGGCTCCGAC	816
Qy	241	TyrSerAspArgProValLeuProProProProProProProProProProProProProPro	260
Db	817	TACTCGGACCGTCCAGTCT	876
Qy	261	SerThrPheThrThrGluLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla	280
Db	877	AGCACTTTCATACGAGGAGTCTAGTACGCCCAATGTTCTCCGAGGCGGCACTTG	936
Qy	281	LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProProSerGlyLysGluVal	300
Db	937	TTAGGACAAAGCGGGTTCGGTTCGTCACAAAGGTGTGTTCCTCTCTCTCTCTCTCTCTCTCT	996
Qy	301	AlaValLysGlnLeuLysValGlySerGlyGlnGlyGlyGlyGlyGlyGlyGlyGlyGlyGly	320
Db	997	GCTGTGAACGAGTTGAAGTTGGAGTGTCTCAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGG	1056
Qy	321	GluThrThrSerArgValHisHisHisHisHisHisHisHisHisHisHisHisHisHisHis	340

Db	1057	GAGATCATCAGCAGAGTTCCACAGGCATCTGGTGTCTCTTGTGGTATTATTCATCGCC	1116
Qy	341	GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu	360
Db	1117	GGTGCCAAAGATTGCTTGTCTATGAGTTTGTCTTAAACAACAATCTCGAGCTTCACCTC	1176
Qy	361	HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysLysLysLysLysLysLys	380
Db	1177	CATGCGAGGAGCGGCTCAATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT	1236
Qy	381	AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysLysLysLysLysLysLysLys	400
Db	1237	GCTAAAGGACTTCTTATCTTCAAGATTGCAATCTTAAATCATTTACCGTGATATC	1296
Qy	401	LysAlaSerAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	420
Db	1297	AAGGCTTCAACATATTGATAGATTTCAGTTTGAAGTTGAAGTTGCTGATTTGGTCTT	1356
Qy	421	AlaLys	440
Db	1357	GCTAAGATTGCTCTGATACAAACACGATGTATCAACGCTGTGATGGGAACCTTTGGG	1416
Qy	441	TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer	460
Db	1417	TACTTGGCTCCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTCA	1476
Qy	461	PheGlyValValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	480
Db	1477	TTTGGCTGTGCTTTTGGAGCTCATTTACTGGAGCTCGACCCGTTGATGCCAACAATGTC	1536
Qy	481	TyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuLeuLeuLeuLeuLeuLeuLeu	500
Db	1537	TATGTAGATGACAGCTTAGTTGACTGGGCGACGACCATTTGCTTAAACCGAGCATCTGAGCA	1596
Qy	501	GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMet	520
Db	1597	GGAGACTTTGAGGCTTTAGCTGATGCAAGATGAATTAATGGGTATGACAGAGAGGAGATG	1656
Qy	521	AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaAlaArgArgProArg	540
Db	1657	GCTCGCATGTTGCTTGTGCTGGCTTGTTCGCCATTCAGCTCGCGGAGACCTCGC	1716
Qy	541	MetSerGlnThrValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly	560
Db	1717	ATGAGCCAGATTGTCGCTGCTTAGAAGGAAATGTATCACTCTCAGATCTTTAACAAGGG	1776
Qy	561	MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer	580
Db	1777	ATGAGACCAAGTCAAGCAATGTATACAGCTCATACGAGGAGGAGCACCAGATTATGACTCG	1836
Qy	581	SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr	600
Db	1837	AGCAGTACATGAAACATGAAGAGTTTAGGAAATGGCACTTGGAACTCAAGAGTAC	1896
Qy	601	AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer	620
Db	1897	AACGCCAGGGTGAGTACAGTAATCCGACGAGTATGAGCTGTACCTCGCTCTGCTTCA	1956
Qy	621	SerSerGluGlyGlnThrThrArgGluMetGlyLysLysLysLysLysLysLysLysLysLys	640
Db	1957	AGCAGCGAGGCGCAACACACACGAAATGGAGATGGGGAAGATTAAAGAGAACCGGTGAG	2016
Qy	641	GlyTyrSerGlyProSerLeu	647
Db	2017	GGTTATAGTGAGCTTCTCTT	2037

RESULT 4

ACF36557  
ID ACF36557 standard; DNA; 1959 BP.

XX ACF36557;

AC ACF36557;

XX 18-DEC-2003 (first entry)  
DT

XX A. thaliana PERK1 protein encoding genomic DNA.  
 XX DE  
 XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;  
 KW wound; pathogen resistance; plant growth; seed production; gene; ds.  
 XX  
 XX Arabidopsis thaliana.  
 OS  
 XX WO2003072763-A1.  
 FN  
 XX 04-SEP-2003.  
 PD  
 XX 28-FEB-2003; 2003WO-CA000274.  
 XX  
 XX 28-FEB-2002; 2002CA-02373903.  
 PR  
 XX 28-FEB-2002; 2002US-00086464.  
 XX  
 XX (GORI/) GORING D.  
 PA (SILV/) SILVA N.  
 PA (HAFF/) HAFFANI Y Z.  
 XX  
 XX Goring D, Silva N, Haffani YZ;  
 FI  
 XX WPI; 2003-712727/67.  
 DR  
 XX Producing a transgenic plant having an increased plant resistance, plant  
 PT growth or seed production comprises transforming a plant with a nucleic  
 PT acid molecule having a Proline-rich Extensin-like Receptor Kinase  
 PT activity.  
 XX  
 XX Disclosure; Fig 21d; 123pp; English.  
 XX  
 CC The invention relates to producing a transgenic plant having increased  
 CC plant height, number of branches, number of seed pods and/or seed  
 CC production compared to a non-transgenic plant, and/or quicker flowering  
 CC or later senescence compared to a non-transgenic plant. The method  
 CC involves transforming a plant with a vector including a Proline-rich  
 CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic  
 CC acid molecule having PERK activity. The method, as well as the PERK  
 CC nucleic acid molecule and polypeptide, are useful in increasing plant  
 CC resistance to wounding and pathogens and in increasing plant growth and  
 CC seed production. The nucleic acid molecule and polypeptide may also be  
 CC used in producing transgenic plants or transgenic host cells. The present  
 CC sequence represents an A. thaliana PERK1 DNA (At3g24550)  
 XX  
 SQ Sequence 1959 BP; 496 A; 514 C; 425 G; 524 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 5 28e-78 Length: 1959  
 Score: 2936.50 Matches: 574  
 Percent Similarity: 90.61% Conservative: 24  
 Best Local Similarity: 86.97% Mismatches: 41  
 Query Match: 85.04% Indels: 21  
 DB: 9 Gaps: 12

US-10-086-464-2 (1-647) x ACP36557 (1-1959)

QY 1 MetSerSerAlaProSerProGlyThr-----GlySerProProSerProProSer 17  
 DB 1 ATGTCCACAGCGCGTCTCCAGGCACCTACTCCATCACCATCTCCACCGTCTCTCCACCA 60  
 QY 18 AsnSerThrThrThrProProAla-----SerAlaProProProThrPro 35  
 DB 61 AACTCGACACACACACACTCTCTCTCCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
 QY 36 SerSerProProPro---ProSerThrThrProSerProProSerSerArgSer 54  
 DB 121 TCTTCT 180  
 QY 55 ThrProSerAlaProProProSerProProThrProProThrProGlySerProProPro 74  
 DB 181 CCTCCT 228

QY 75 LeuProGlnProSerProProAlaProThrThrThrProGlySerProProAlaProValThr 94  
 DB 229 CTTCTCTCAACCTTCCCGCTCGCTCCCATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 285  
 QY 95 ProProThr---ArgAsnProProProSerValProGlyProProProSerAsnProSerArg 113  
 DB 286 CCCTCAAAACCTCGAAGCCCTCCATCTCTTAACCAAGGACCAACCAACACCTCCCTCA--- 342  
 QY 114 GluGlyGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131  
 DB 343 ---GGATCTACTCTAGAACTCCATCAACACTAAACCGTCCGCGCGCTCTGATTTCTTCC 399  
 QY 132 AspGlyLeuSerThrGlyValValGlyValGlyLeuLeuLeuVal 151  
 DB 400 GATGATTTGCTACCGGAGTTGTGTAGGAATCCCATTTGGAGGAGTTCCTATTCTTTGTT 459  
 QY 152 IleValThrLeuLeuCysLeuLeuCysLysLysLysArgArgArg-----AspGluGlu 169  
 DB 460 ATACTGACTCTGATTTGTCTTCTCTGTAAGAAGAAACGAAGAAGACACACGACGATGAA 519  
 QY 170 AspAlaTyrTyrValProProProProGlyProGlyProLysAlaGlyGlyProTyrGly 189  
 DB 520 GCTGCTTACTATGTTCT 579  
 QY 190 GlyGlnGlnGlnGlnTTPArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208  
 DB 580 GGTCAACAACAGTATTGGCAACAACAACACGCTACCGCGCTCAGATAATCATGTAGTG 639  
 QY 209 ThrSerLeuProProProProLysAlaProSerProProArgGlnProProProProPro 228  
 DB 640 AGTTCATG---CCACCACCTAAGCCTCCATCTCCACCACGAAACCTCTCTCGCCACCT 696  
 QY 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247  
 DB 697 CCACCACGACCATTCATGAGTAGCAGTGGTGGTCTCTGACTATTTCGGATCTTCCGGTCTT 756  
 QY 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267  
 DB 757 CCTCACCATCTCCAGGCGTTGTGTAGCTTTCTTAAAGCATTCTTCACTTATGAGGAG 816  
 QY 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287  
 DB 817 TTGTCGAGAGCTACTAATGGCTCTCTGAGGCTAATTTGTTAGGACAAGGAGGTTTGGT 876  
 QY 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307  
 DB 877 TATGTGCATAAAGGTATATTGCTTAGTGGGAAAGATTGCTGTGAAACAGTTGAAAGCT 936  
 QY 308 GlySerGlyGlnGlyGluValGluPheGlnAlaGluValGluLeuLeuLeuSerArgValHis 327  
 DB 937 GGTAGTGGTCAGGAGAGAGAGAGAGTTTCAGGCTGAGGTTGAGATCATTAGCAGAGTTTCT 996  
 QY 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347  
 DB 997 CACAGGCATTGTTGTTCTCTTATTGTTATTGTTATGTCGGCGGTGTTCAAAAGATTACTTCTC 1056  
 QY 348 TyrGluPheValProAsnAsnLeuGluLeuHisIleuHisGlyGluGlyArgProThr 367  
 DB 1057 TATGAGTTGTTTCCAAACACAATCTTGAGTTTTCACCTCCATGTTGAGGAGCGCCCTACG 1116  
 QY 368 MetGluTTPSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLysSerTyrLeu 387  
 DB 1117 ATGGAATGGAGTACTAGATTGGAATTTGCTCTTGGATCTGCTTAAGGACTTTTCATATCTT 1176  
 QY 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 407  
 DB 1177 CATGAAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGCGCTCAACATATTGATT 1236  
 QY 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427  
 DB 1237 GATTTCAAATTTGAAGCTTAAGGTTGCTGACTTTGGTCTTTCGCAAGATTGCTTCTGATACA 1296  
 QY 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447

```
Db      ||||| 1297 AACACTCATGTATCTACACGGGTGATGGAAACCTTTGGGTATTTGGCTCCGGAATAGCT 1356
Qy      ||||| 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
Db      ||||| 1357 GCAAGTGAAGAGCTCACAGAAAGTCTGACGTTTCTCATTTGGCGTTGTACTTTTGAA 1416
Qy      ||||| 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspSerLeuVal 487
Db      ||||| 1417 CTTATTACTGGAGGCGCTGTTGATCGCAACATGTCATGTAGATGACAGCTTAGTT 1476
Qy      ||||| 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAla 507
Db      ||||| 1477 GACTGGGCACCACTTCTTAACCGAGCATCTGAGGAAGGAGATTTTGAGGGTTTGCT 1536
Qy      ||||| 508 AspAlaLysMetAsnAenGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAla 527
Db      ||||| 1537 GATTCAAGATGGGTATGATGATGACAGAGGAGATGGCTCGCATGTTGCTTGGCT 1596
Qy      ||||| 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547
Db      ||||| 1597 GCGGCTTGTTGCCATTACGTCGCGCAGACCTCGCATGAGCCAGATAGTACGGCG 1656
Qy      ||||| 548 LeuGluGlyAenValSerLeuSerAspLeuAenGluGlyMetArgProGlyGlnSerAen 567
Db      ||||| 1657 TTAGAAGAAATGATCGCTGCTCTCATCTTAACGAAGGAGTACAGCCGGTTCACAGCAAC 1716
Qy      ||||| 568 ValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAenGluAspMet 587
Db      ||||| 1717 GTATACAGCTCATATGAGGAAGGACAGACTATGACAGAGCCCAATACAGCAGCATG 1776
Qy      ||||| 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAenAlaThrGlyGluTyrSer 607
Db      ||||| 1777 ATAAAGTTAGGAATAGGCTCTTGAACTCAAGNATACGCACAAACCGGCGAGTACAGT 1836
Qy      ||||| 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThr 627
Db      ||||| 1837 AATCCAACCACTAGTACGGAAGTACCGGCTGTCGTTCAAGCAGTGAAGGTCAAGCCACA 1896
Qy      ||||| 628 ArgGluMetGluMetGlyLysLeuLysArgThrGlnGlyTyrSerGlyProSerLeu 647
Db      ||||| 1897 CGAGAAATGGAGATGGGAAGATTAAGAAACCGGTCAAGGTTATAGTGGACCTCTCTT 1956

RESULT 5
ID AAC36968 standard; DNA; 2324 BP.
XX
AC AAC36968;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15710.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX
99US-0128714P.
99US-0129845P.
99US-0130077P.
99US-0130449P.
99US-0130510P.
99US-0130891P.
99US-0131449P.
99US-0132048P.
99US-0132407P.
99US-0132484P.
99US-0132485P.
99US-0132486P.
99US-0132487P.
99US-0132863P.
99US-0134256P.
99US-0134218P.
99US-0134219P.
99US-0134221P.
99US-0134370P.
99US-0134768P.
99US-0134941P.
99US-0135124P.
99US-0135353P.
99US-0135629P.
99US-0136021P.
99US-0136392P.
99US-0136782P.
99US-0137222P.
99US-0137528P.
99US-0137724P.
99US-0138094P.
99US-0138540P.
99US-0138847P.
99US-0139119P.
99US-0139452P.
99US-0139453P.
99US-0139492P.
99US-0139454P.
99US-0139455P.
99US-0139456P.
99US-0139457P.
99US-0139458P.
99US-0139459P.
99US-0139460P.
99US-0139461P.
99US-0139462P.
99US-0139463P.
99US-0139750P.
99US-0139763P.
99US-0139817P.
99US-0139899P.
99US-0140353P.
99US-0140354P.
99US-0140695P.
99US-0140823P.
99US-0140991P.
99US-0141287P.
99US-0141842P.
99US-0142154P.
99US-0142055P.
99US-0142390P.
99US-0142803P.
99US-0142920P.
99US-0142977P.
99US-0143542P.
99US-0143624P.
99US-0144005P.
99US-0144085P.
99US-0144086P.
99US-0144325P.
99US-0144331P.
99US-0144332P.
```

PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.
PR 26-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161405P.
PR 27-JUL-1999;	99US-0145913P.	PR 25-OCT-1999;	99US-0161406P.
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161359P.
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161360P.
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161361P.
PR 02-AUG-1999;	99US-0146386P.	PR 28-OCT-1999;	99US-0161920P.
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161922P.
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161933P.
PR 03-AUG-1999;	99US-0147038P.	PR 29-OCT-1999;	99US-0162142P.
PR 04-AUG-1999;	99US-0147204P.		
PR 05-AUG-1999;	99US-0147302P.	Alignment Scores:	
PR 05-AUG-1999;	99US-0147192P.	Pred. No.:	1.88e-76
PR 05-AUG-1999;	99US-0147260P.	Score:	2885.00
PR 06-AUG-1999;	99US-0147303P.	Percent Similarity:	90.30%
PR 06-AUG-1999;	99US-0147416P.	Best Local Similarity:	86.36%
PR 09-AUG-1999;	99US-0147493P.	Query Match:	83.55%
PR 09-AUG-1999;	99US-0147935P.	DB:	3
PR 10-AUG-1999;	99US-0148171P.		
PR 11-AUG-1999;	99US-0148319P.	US-10-086-464-2 (1-647) x AAC36968 (1-2324)	
PR 12-AUG-1999;	99US-0148341P.		
PR 13-AUG-1999;	99US-0148563P.		
PR 13-AUG-1999;	99US-0148684P.	Qy 1 MetSerSerAlaProSerProGlyThr-----GlySerProSerProSerProSer 17	
PR 16-AUG-1999;	99US-0149368P.	Db 105 ATGTCCACAGCGCGTCTCCAGGCACCTACTCCATCAGCATCTCCACCGTCTCTCCCCACA 164	
PR 17-AUG-1999;	99US-0149175P.		
PR 18-AUG-1999;	99US-0149426P.	Qy 18 AsnSerThrThrThrProProAla-----SerAlaProProProProProPro 35	
PR 20-AUG-1999;	99US-0149723P.	Db 165 AACTCGACACCCACTCTCTCTCCAGCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 224	
PR 20-AUG-1999;	99US-0149902P.		
PR 23-AUG-1999;	99US-0149902P.	Qy 36 SerSerProProPro-----ProSerThrIleProThrSerProProProSerSerArgSer 54	
PR 25-AUG-1999;	99US-0150566P.	Db 225 TCTTCTT 284	
PR 26-AUG-1999;	99US-0150884P.		
PR 27-AUG-1999;	99US-0151065P.	Qy 55 ThrProSerAlaProProProProSerProProProSerProProThrProGlySerProProPro 74	
PR 27-AUG-1999;	99US-0151066P.	Db 285 CCTCTTCT 332	
PR 27-AUG-1999;	99US-0151080P.		
PR 30-AUG-1999;	99US-0151303P.	Qy 75 LeuProGlnProSerProProAlaProThrThrProGlySerProProAlaProValThr 94	
PR 31-AUG-1999;	99US-0151438P.	Db 333 CTTCTCAACCTTCTCCATCCGCTCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 389	
PR 01-SEP-1999;	99US-0151930P.		
PR 07-SEP-1999;	99US-0152363P.	Qy 95 ProProThr---ArgAsnProProProSerValProGlyProProProSerAsnProSerArg 113	
PR 10-SEP-1999;	99US-0153070P.	Db 390 CCCTCAACCTCGAAGCCCTCCATCTCTCAACCAAGGACCAACCAACACCTCCCTCA--- 446	
PR 13-SEP-1999;	99US-0153758P.		
PR 15-SEP-1999;	99US-0154018P.	Qy 114 GluGlySerProArgProProSerSer-----ProSerProProSerProSerSer 131	
PR 16-SEP-1999;	99US-0154039P.	Db 447 ---GGATCTACTCTAGAACTCTCAACAGCTTAACCGTGAACCGTCCGCCACCGTCTTGATCTTCC 503	
PR 20-SEP-1999;	99US-0154779P.		
PR 22-SEP-1999;	99US-0155139P.	Qy 132 AspGlyLeuSerThrGlyValValGlyValGlyValGlyValGlyValGlyValGlyValGlyVal 151	
PR 23-SEP-1999;	99US-0155486P.	Db 504 GATGATTTGTTACCGGAGTTGGTAGGATTCGCCATTCGCCATTCGCCATTCGCCATTCGTTGTT 563	
PR 24-SEP-1999;	99US-0155655P.		
PR 28-SEP-1999;	99US-0156458P.	Qy 152 IleValThrLeuLeuCysLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 169	
PR 29-SEP-1999;	99US-0156596P.	Db 564 ATACTGACTCTGATTTTGTCTTCTCTGTAAGAAAGAAACGAAAGAAAGACACGACGATGAA 623	
PR 04-OCT-1999;	99US-0157117P.		
PR 05-OCT-1999;	99US-0157753P.		
PR 06-OCT-1999;	99US-0157865P.		
PR 07-OCT-1999;	99US-0158029P.		
PR 08-OCT-1999;	99US-0158232P.		
PR 12-OCT-1999;	99US-0158369P.		
PR 13-OCT-1999;	99US-0159293P.		
PR 13-OCT-1999;	99US-0159294P.		

Qy 170 AspAlaTyrTyrValProProProProProProGlyProLysAlaGlyGlyProTyrGly 189  
Db 624 GCTGTTACTATGTT---CCTCCTCCTCCTTCTGCTCCAAAGCTGGAGGACCTTACGGT 680  
Qy 190 GlyGlnGlnGlnTTPArgGlnGlnAsnAlaThrProProSerAsp---HisValVal 208  
Db 681 GGTCAACAACAGTATGTGGCAACAACAACCGCTACCGCGCTCAGATAATCATGTAGTG 740  
Qy 209 ThrSerLeuProProProProProProProProProProProProProProProPro 228  
Db 741 ACGTCAATG---CACACACCTTAAGCTTCATCTCCACCACCAAACTCTCCGCCACCT 797  
Qy 229 ProProPro---PheMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeu 247  
Db 798 CCACCACGACATTCATGAGTAGCAGTGGTTCCTGACTATTCGGATCTTCCGGTTCCT 857  
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267  
Db 858 CCTCCACCATCTCCAGGCTTGTGTAGGCTTTTCAAAAAGCACTTTCACCTATATGAGGAG 917  
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGly 287  
Db 918 TTGTCAAGAGCTACTAATGGCTTCTCTAGAGCTAATTTGTAGGACAAAGGAGGTTTGGT 977  
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307  
Db 978 TATGTGCATAAAGTATATTCCTAGTGGGAAAGAGTTGCTGTGAAACAGTTCAAAGCT 1037  
Qy 308 GlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLysLeuSerArgValHis 327  
Db 1038 GGGAGTGTCTAGGAGAGAGAGAGTTCAGGCTGAGGTTGAGATCATTTAGCAGAGTTTCAT 1097  
Qy 328 HisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuVal 347  
Db 1098 CACAGGATTTGGTTCCTTATTGTTATTGATGCTGATGTTCAAGATTGCTTGTGTC 1157  
Qy 348 TyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlyArgProThr 367  
Db 1158 TATGAGTTTGTCTTCAACAACATCTTGAGTTTCACCTCCATGTTAAGGGAAGCCTTACA 1217  
Qy 368 MetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeu 387  
Db 1218 ATGGAATGGAGTACTAGATTGAAGATTGCTCTTGATCTGCTAAAGGACCTTTCATATCTT 1277  
Qy 388 HisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIle 407  
Db 1278 CATGAGATTGCAATCCGAAATCATTCACCGTGATATTAAAGGCTCAACATATTGATT 1337  
Qy 408 AspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThr 427  
Db 1338 GATTTCAAAATTTGAAGCTAAGTTGCTGACTTTGCTGCTGCGAGATTGCTTCTGTATACA 1397  
Qy 428 AsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447  
Db 1398 AACACTCATGTATCTACACGCTGATGGGAACCTTTGGGTATTTGGCTCCAGATAATGCT 1457  
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467  
Db 1458 GCAAGCGGAAGCTCACAGAAAAGTCTGACGTTTCTCATTTGTTGTTGTTGTTGGAG 1517  
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspSerLeuVal 487  
Db 1518 CTTATACCGGGAGCGCCCTTGTGATGCAACAACATGCTATGTAGATGACAGCTTATGTT 1577  
Qy 488 AspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAla 507  
Db 1578 GACTGGGCACGACCTTGTCTTAACCGAGCATCTCAGGAAGGAGATTTTGTGTTGGCT 1637  
Qy 508 AspAlaLysMetAsnAsnGlyTyrAspArgGluLysMetAlaArgMetValAlaCysAla 527  
Db 1638 GATTCAAGATGGGTAAATAGTATAG-AGAGAGAGAGATGCTCCATGTTGCTTGGCT 1696  
Qy 528 AlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAla 547

Db 1697 GCAGCTTGTGTTGCCATTTCAGCTCGCGCAGACCTCGCATGAGCCAGATAGTACGGCG 1756  
Qy 548 LeuGluGlyAsnValSerLeuSerAspLeuAsnGlyMetArgProGlyGlnSerAsn 567  
Db 1757 TTAGAAGGAATGTATCGCTGCTGTATCTTAACGAAGGGATGAGACCGGGTCA CAGCAAC 1816  
Qy 568 ValTyrSerSerTyrGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMet 587  
Db 1817 GTATACAGCTCATATGGAGGAGCAGCTGACTATGACACAGGCCAATACACAGATGACATG 1876  
Qy 588 LysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSer 607  
Db 1877 AAGAAGTTTAGGAAATGGCTCTTGGAACTCAAGATACGCAACACCGCGAGTACAGT 1936  
Qy 608 AsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThr 627  
Db 1937 AATCAACACAGTACTACGAGCTGTACCCGCTCTGTTCAAGCAGTGAAGGTCAAGCCACA 1996  
Qy 628 ArgGluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647  
Db 1997 CGAGAAATGGAGATGGGAAGATTAGAAAACCGTCAAGGTTATAGTGGACCTCTCTT 2056  
RESULT 6  
ACF36554  
ID ACF36554 standard; DNA; 1812 BP.  
XX  
AC ACF36554;  
DT 18-DEC-2003 (first entry)  
XX  
DE O. sativa PERK protein encoding genomic DNA.  
XX  
KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;  
KW wound; pathogen resistance; plant growth; seed production; rice; gene;  
KW ds.  
OS Oryza sativa.  
XX  
PN W02003072763-A1.  
XX  
PD 04-SEP-2003.  
XX  
PF 28-FEB-2003; 2003WO-CA000274.  
XX  
PR 28-FEB-2002; 2002CA-02373903.  
PR 28-FEB-2002; 2002US-00086464.  
XX  
(GORI/) GORING D.  
(SILV/) SILVA N.  
(HAFF/) HAFFANI Y Z.  
PI Goring D, Silva N, Haffani YZ;  
XX  
DR WPI; 2003-712727/67.  
XX  
PT Producing a transgenic plant having an increased plant resistance, plant  
PT growth or seed production comprises transforming a plant with a nucleic  
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase  
PT activity.  
XX  
PS Disclosure; Fig 21a; 123pp; English.  
XX  
CC The invention relates to producing a transgenic plant having increased  
CC plant height, number of branches, number of seed pods and/or seed  
CC production compared to a non-transgenic plant, and/or quicker flowering  
CC or later senescence compared to a non-transgenic plant. The method  
CC involves transforming a plant with a vector including a proline-rich  
CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic  
CC acid molecule having PERK activity. The method, as well as the PERK  
CC nucleic acid molecule and polypeptide, are useful in increasing plant  
CC resistance to wounding and pathogens and in increasing plant growth and  
CC seed production. The nucleic acid molecule and polypeptide may also be

CC used in producing transgenic plants or transgenic host cells. The present  
CC sequence represents an *O. sativa* PERK protein encoding genomic DNA (TIGR  
XX Accession No. TC102111)

SQ Sequence 1812 BP; 409 A; 487 C; 512 G; 404 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	1.6e-45	Length:	1812
Score:	1821.50	Matches:	370
Percent Similarity:	71.62%	Conservative:	69
Best Local Similarity:	60.36%	Mismatches:	119
Query Match:	52.75%	Indels:	55
DB:	9	Gaps:	15

US-10-086-464-2 (1-647) x ACF36554 (1-1812)

Qy	54	SerThrProSerAlaProProProSerProProThrProSerThrProGlySerProPro	73
Db	4	TCGTGCGCGTCTCGCGCGCGCGCGTAAACAGAGC-----GGACACCGCCT	51
Qy	74	ProLeuProGlnProSer-----ProProAlaProThrProGlySerProProAla	91
Db	52	CCGGCTAACACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	111
Qy	92	Pro-----ValThrProProThrArgAsnProProProSerValProGlyProProSer	109
Db	112	CCGGCGCTCTTTGCTCGCGCGCTCGCTGACG---CCTCCCGCAGCGCGCATCGCTCCGTCG	168
Qy	110	AsnProSerArgGluGlyGlySerProArgProProSerProSerProSerProSer	128
Db	169	ACC-----ACGCCCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	201
Qy	129	ProSerSerAspGlyLeuSerThrGlyValValValGlyValAlaIleGlyValAla	148
Db	202	CGCTCGAGCTCGGGCTGACGACCGCGCGTGTGTGGGATCGTGGTGGCGCGCTCGTC	261
Qy	149	LeuLeuValIleValThr-----LeuIleCysLeuLeuCysLysLysLysLysLysLys	166
Db	262	GTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	321
Qy	167	AspGluGluAspAla-----TyrTyr-----Val	174
Db	322	CACACACACCTCCCG	381
Qy	175	ProProProProProProProGlyProGlyProGlyProGlyProGlyProGlyProGly	194
Db	382	CCG	441
Qy	195	TrpArgGlnGlnAsnAlaThrProProSerAspHisValValValThrSerLeuProPro	214
Db	442	TGGCAGCAAAATGGTCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCTCC	483
Qy	215	ProLysAlaProProProProProProProProProProProProProProPheMetSer	234
Db	484	AAGAAGGTGCTTTCGATCTCTCCCTCCACCTCCACCTCCACCTCCACCTCCACCT	543
Qy	235	SerSerGlyGlySerAspTyrSer-----AspArgProValLeuProProPro	250
Db	544	GGTGTCTCTGCTCAAAATTAATCTCCGCTGGCGGCGGCGGCGGCGGCGGCGGCGG	594
Qy	251	SerProGlyLeuValLeuGlyPheSerLysSerThrPheThrThrGluGluLeuAlaArg	270
Db	595	TCACCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	654
Qy	271	AlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlnGlnGlyPheGlyTyrValHis	290
Db	655	CGGACCGGATGATCTCCGATCTAACTCTCGACAGCGCGCGGCTTTGGGTATGTTAC	714
Qy	291	LysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLeuValGlySerGly	310
Db	715	AAAGGAGTTCTGCCGAATGGGACAGAGGTTCTGTGAAGCAATTTGAGAGATGGGAGT	774
Qy	311	GlnGlyGluArgGluPheGlnAlaGluValGluIleLeuSerArgValHisHisArgHis	330

Db	775	CAGGAGAGCGTGAATTTTCAGCGGAGGTTGAGATTATCAGCGGATCATCAACAGCAT	834
Qy	331	LeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPhe	350
Db	835	CTTGTAACATTTGGTTGTTTATTCATTTCTGGAGGGAAGAGTTGCTTGTCTATGAGTAT	894
Qy	351	ValProAsnAsnLeuLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrp	370
Db	895	GTTCCTCAAAACACATTTGGAGCTCCACTTGGGAGAGCGGCCCAACCATCGAATGG	954
Qy	371	SerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAsp	390
Db	955	CCTACAAAGATTACCATTTGCTGGGCTGCGAAGGATGGCATATCTTCATGAAGAT	1014
Qy	391	CysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLys	410
Db	1015	TGCCATCTTAAGATCATTTACCGGTGACATAAAGTCAGCAAAATAATCTTCTTGACGCG	1074
Qy	411	PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis	430
Db	1075	TTTGAGGCAAAAGGTGGCAGATTTTCGACTTTCGAAAATTGACTTCTGTATAATAACAC	1134
Qy	431	ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly	450
Db	1135	GTTCCTCAACAGAGTAATGGGCACATTTGGGTACCTTGACACAGATATGCGTCTCTGGC	1194
Qy	451	LysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThr	470
Db	1195	CAGCTAACTGAGAAATCAGATGCTTTTCTTCGGAGTAATGCTTCTTGAGCTAATAACT	1254
Qy	471	GlyArgArgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrpAla	490
Db	1255	GGCGCGCGCTCTGTGAGATCGAACCAATCAAGATGATGACAGCTTGGTGTGATGGGCA	1314
Qy	491	ArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLys	510
Db	1315	AGGCTTTTAATGTCGAGCATCAGATGATGGCAATATGATGCTTTAGTGGATCCACGG	1374
Qy	511	MetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCys	530
Db	1375	CTAGGACAGGAGTACAATGGTAACGAGATGCGGAGGATGATTGCTTGTGCAAGCTGTGT	1434
Qy	531	ValArgHisSerAlaArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGly	550
Db	1435	GTACGCCATTTGCAACGCCGACCGCCACGATGAGCGAGTGTTCGGGCGCTTGGAAAGC	1494
Qy	551	AsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSer	570
Db	1495	GATGTGTCATTTGGATGATCTGAACCAAGGTTGCTCTCTGGCCATAGCCGATTTCTTGG	1554
Qy	571	SerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPhe	590
Db	1555	TCATAC--AACACCAACGAGTACGACACGCGGCCCATTTACAATGAAGACCTGAAGAAAT	1611
Qy	591	ArgLysMetAlaLeuGlyThrGlnGlyTyrAsnAlaThrGlyGlyTyrSerAsnProThr	610
Db	1612	AGGAAGATGGCATTTGGCAGTGGCAATCTA-----GAGAGCAGCGACCAACACG	1659
Qy	611	SerAspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGlnThrThrArg-----	628
Db	1660	CAGCCCAACCGAGTTGCTCCGAATCGATCGGTATCAATGGGGGATGCGCGTCAAATACCG	1719
Qy	629	-----GluMetGluMetGlyLysIleLysArgThrGly	639
Db	1720	GAGACCGGAGATGGGAGTGGGAGTTTCAAGAAAGATGGC	1758

RESULT 7

ACF36555

ID ACF36555 standard; DNA; 1515 BP.

XX

AC ACF36555;

XX



Db 1276 GGCTACAGTGGTGCTACAGCTCAGGATACAGCGAGCCTCG 1317

RESULT 8

ACF36556

XX ACF36556 standard; DNA; 1488 BP.

XX

XX ACF36556;

XX

DT 18-DEC-2003 (first entry)

XX

DE G. max PERK partial EST sequence.

XX

XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;

KW wound; pathogen resistance; plant growth; seed production; soybean; gene;

KW ds.

XX

OS Glycine max.

XX

XX WO2003072763-A1.

XX

PD 04-SEP-2003.

XX

PF 28-FEB-2003; 2003WO-CA000274.

XX

PR 28-FEB-2002; 2002CA-02373903.

PR

PR 28-FEB-2002; 2002US-00086464.

XX

XX (GORI/) GORING D.

PA (SILV/) SILVA N.

PA (HAFF/) HAFFANI Y Z.

XX

XX Goring D, Silva N, Haffani YZ;

PI

XX WPI; 2003-712727/67.

DR

XX

XX

PT Producing a transgenic plant having an increased plant resistance, plant growth or seed production comprises transforming a plant with a nucleic acid molecule having a Proline-rich Extensin-like Receptor Kinase activity.

PT

XX

PS Disclosure; Fig 21c; 123pp; English.

XX

CC The invention relates to producing a transgenic plant having increased plant height, number of branches, number of seed pods and/or seed production compared to a non-transgenic plant, and/or quicker flowering or later senescence compared to a non-transgenic plant. The method involves transforming a plant with a vector including a Proline-rich Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic acid molecule having PERK activity. The method, as well as the PERK nucleic acid molecule and polypeptide, are useful in increasing plant resistance to wounding and pathogens and in increasing plant growth and seed production. The nucleic acid molecule and polypeptide may also be used in producing transgenic plants or transgenic host cells. The present CC sequence represents a G. max PERK EST (expressed sequence tag) partial CC sequence (TIGR Accession No. TC123698)

XX

SQ Sequence 1488 BP; 412 A; 291 C; 348 G; 437 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3,246-40	Length:	1488
Score:	1637.50	Matches:	312
Percent Similarity:	89.67%	Conservative:	44
Best Local Similarity:	78.59%	Mismatches:	36
Query Match:	47.42%	Indels:	5
DB:	9	Gaps:	3

US-10-086-464-2 (1-647) x ACF36556 (1-1488)

Qy 253 GlyLeuValLeuGlyPheSerIysSerThrPheThrTyrgluLeuAlaAraGalaThr 272

Db 3 GGAATTCATTGGGGTTCTCTAAGAGCACATTCACGTATGAGGAGTTGGCACGGCAACT 62

Qy 273 AsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrgluValHisIysGly 292

Db 63 GATGGCTTCTCTGATGCCAACCTCTTGGACAAGGAGATTGGATATGTGCACAGAGA 122

Qy 293 ValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGly 312

Db 123 ATTCTTCCCAACGGCAGGAGGTGGCAGTGAAGCAATTGAAGGCTGGAAGCGGCAAGGG 182

Qy 313 GluArgGluPheGlnAlaGluValGluLeuLeuSerArgValHisHisArgHisLeuVal 332

Db 183 GAGCGTGAATTTCCAAGCTGAAGTTGAGATAATTAGCCGTGTCCATCAACAGCATCTTGT 242

Qy 333 SerLeuValGlyTyrcysilleAlaGlyAlaLysArgLeuLeuValTyrgluPheValPro 352

Db 243 TCTTTGGTTGGATCTGATCATCTGGGTCCAGAGGCTGCTTGTATTGAATTTGTTCCC 302

Qy 353 AsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSerThr 372

Db 303 AACAAACATTTGGAAATTCATTTGCATGGAAAAGGAGACCTACCATGGATTGCCACACA 362

Qy 373 ArgLeuLysilleAlaLeuGlySerAlaLysGlyLeuSerTyrgluHisGluAspCysAsn 392

Db 363 AGACTAAGAATTGCTTTAGGATCTGCTAAGGAGCTGGCGTATCTTCATGAAGATTGTCAT 422

Qy 393 ProLysilleHisArgAspileLysAlaSerAsnilleLeuLysAspPheLysPheGlu 412

Db 423 CCTAAGATCATCCATCTGATATCAAGCTGCCAACATCTCTCTGGATTTTAAGTTGAA 482

Qy 413 AlaLysValAlaAspPheGlyLeuAlaLysilleAlaSerAspThrAsnThrHisValSer 432

Db 483 GCAAAGGTTGCAGATTTCCGGTCTTGCAAGTTTCTTCTGATGTCAATACTCATGTTTCT 542

Qy 433 ThrArgValMetGlyThrPheGlyTyrgluAlaProGluTyrgluAlaAlaSerGlyLysLeu 452

Db 543 ACTCGAGTGATGGGGACTTTTGGGTATTTGGCTCCAGATATGCTTCTAGTGGAAAACTG 602

Qy 453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuLysThrGlyArg 472

Db 603 ACACACAATCAGATGTTTCTCTACGGAGTCATGCTCTCGAGTTAATAACCGGACGA 662

Qy 473 ArgProValAspAlaAsnValTyrgluValAspAspSerLeuValAspTrpAlaArgPro 492

Db 663 CGGCCCGTCGATAAAATCAAACTTTCATGGAGGATAGTTTGGTAGACTGGGCTAGGCCT 722

Qy 493 LeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAspAlaLysMetAsn 512

Db 723 TTGTCTCACAGAGCTTTGGAAAGAGGATGATTTGATTTATTATGACCCCAAGGCTCCAG 782

Qy 513 AsnGlyTyrgluArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532

Db 783 AATGACTATGATCTCTAATGATGAGTGGCAGATGTTGGCTTCTGCTGGGCTTGACATAGT 842

Qy 533 HisSerAlaArgArgArgProArgMetSerGlnilleValArgAlaLeuGluGlyAsnVal 552

Db 843 CATTCGGCAAAAGCTCGACCAAGGATGAGCAGGTTGTCCGCGCTCTCGAAGAGAGATGTC 902

Qy 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrgluSerSerTyrg 572

Db 903 TCTCTAGCAGATCTGAATGAAGGAATGAAGCCTGGACACAGCATATGATAGTTTCTCAT 962

Qy 573 GlyGlySerThrAspTyrgluSerSerGlnTyrgluAsnGluAspMetLysLysPheArgLys 592

Db 963 ---GAAAGCTCAGATTATGACACTGCACAGTACAAAGGAAGACATGAAAAGTTTCAGGAAA 1019

Qy 593 MetAlaLeuGlyThrGlnGluTyrgluAsnAlaThrGlyGluTyrgluSerAsnProThrSerAsp 612

Db 1020 ATGGCATTTGGGAACCTCAGGAGTATGGTGAAGAGTGTAGTACAGTCCGCGCTACAAAGTGAG 1079

Qy 613 TyrglyLeuTyrgluProSerGlySerSerGluGly-----GlnThrThrArgGlu 629

Db 1080 TATGGTTTAAACCCATCAGGCTCAAGTAGTGAACAGACAGCCGCCCAACCAACCAAGGGAA 1139

Qy 630 MetGluMetGlyLysilleLysArgThrGlyGlnGlyTyrgluTyrgluProSer 646



QY 368 etGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuH 388  
 Db 1302 TGGAGTTCTCCACTAGTTGCGTATCGCTAGGTGCTCGCAAGGACTCGCTTACCTTC 1361  
 QY 388 IeGluAspCys----- 391  
 Db 1362 ACGAAGACTGTAAGTTTTAAACATTCCACCTTCATCTTTTCTTAACCAAGTTGCATAAAAC 1421  
 QY 392 -----Asn 392  
 Db 1422 AGAGAAAGCTCTGCTCTGACTAGTGTATCTTTTGGCTGAGAAAAATGGTGCAGGCCAT 1481  
 QY 393 ProLysIleIleHisArgAspIleLysAlaSerAsenIleLeuIleAspPheLysPheGlu 412  
 Db 1482 CCTCGGATCAITTCACCGGACATCAAGTCTGCAATATTTCTTTGGACTTCAACTTTGAT 1541  
 QY 413 Ala----- 413  
 Db 1542 GCTATGGTGATAAACTAGTAGCTTCATTCATCTACGGTTTTTTGTTAAGACTACATTGA 1601  
 QY 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAs 426  
 Db 1602 TGACATTTTGCAATTGTTTATTCAGTGGCTGATTTTGGATTAGTAACTAATCACTGA 1661  
 QY 426 pThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440  
 Db 1662 TAAACAACACTCATGTATCTACTCGTGTGTATGGAACTTTTCGGGTAAAGCGTTTACCGTAT 1721  
 QY 440 ----- 440  
 Db 1722 GATAAGATTGTCGTGACACTCAAGAAACATAACCTTTGTAGACTAATCTACTTTGCTTC 1781  
 QY 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrG 454  
 Db 1782 TTTCCACAAACATGTGTAGATATCTAGCTCCAGATATGCTTCAAGCGTAAATTAACCG 1841  
 QY 454 LuLysSerAspValPheSerPheGlyValValLeuLeuLeuIleThrGlyArgArgP 474  
 Db 1842 AGAAATCCGATGTTTTCTCTTACGGAGTTATGTTATTGGAACTTATAACTGGMAACGAC 1901  
 QY 474 roValAspAlaAsnValTyrValAspAspSerLeuValAspTrp----- 489  
 Db 1902 CGGTTGAT---AATGATCATCCATGGAGCACCTTAGTAGATTGGTATTCATGCAATG 1958  
 QY 489 ----- 489  
 Db 1959 TAACATATGATCGGTATATATGTTTTTCGCTTTTCGCTACTAATGATCATGAATA 2018  
 QY 490 ---AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508  
 Db 2019 CAGGCTCGGCTCTTTATGCTCGCGCTAGAGATGGAACCTTTAATGAGCTCGCAGAT 2078  
 QY 509 AlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAla 528  
 Db 2079 GCGAGGCTGAAGGCAACTACAAACCGCAAGAAATGGCTCGAATGGTGACTTGTCCGCT 2138  
 QY 529 AlaCysValArgHisSerAlaArgArgProArgMetSer----- 542  
 Db 2139 GCTAGCATTCGTCATTCGGGGGTAAACGCTCCAAAGATGAG-CCAGGTGAATCAAAATTA 2197  
 QY 543 -----GlnIleVal 545  
 Db 2198 TAACTAAAGTCTATTTTTTGTGTCAGAGAATAACAAACAATGTTGGTTTTTCAGATAGTA 2257  
 QY 546 ArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGln 565  
 Db 2258 AGAGCGTTTAGAGAGAGAGATGCTTATGATGCTTTAAACGAAGGTGTGAAGCAGACAC 2317  
 QY 566 SerAsnValTyrSerSerTyrGlySerThrArgPheAspSerSerGlnTyrAsnGlu 585  
 Db 2318 ACTAAGCTTTACGGGTCTATTGGGAGCAAGCTCGGATTATAGTCAGACATCTTACAAATGCA 2377

QY 586 AspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr-----AsnAla 602  
 Db 2378 GACATGAAGAAATTCAGACAGATAGCTTTGTCGAGCAAGAAATTCACAGTCAGTACTGT 2437  
 QY 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGly 614  
 Db 2438 GAAGGACATCTAGTAATGATTCTAGAGATATGGGA 2473  
 RESULT 10  
 ACF36551  
 ID ACF36551 standard; DNA; 2520 BP.  
 XX AC ACF36551;  
 XX AC ACF36551;  
 DT 18-DEC-2003 (first entry)  
 XX Arabidopsis PERK1 receptor related protein encoding DNA.  
 DE Arabidopsis PERK1 receptor related protein encoding DNA.  
 XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;  
 KW wound; pathogen resistance; plant growth; seed production; gene; ds.  
 XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 PN WO2003072763-A1.  
 XX 04-SEP-2003.  
 PD 28-FEB-2003; 2003WO-CA000274.  
 XX 28-FEB-2002; 2002CA-02373903.  
 PF 28-FEB-2002; 2002US-00086464.  
 PR 28-FEB-2002; 2002US-00086464.  
 XX (GORI/) GORING D.  
 PA (SILV/) SILVA N.  
 PA (HAFF/) HAFFANI Y Z.  
 XX Goring D, Silva N, Haffani YZ;  
 XX WPI; 2003-712727/67.  
 DR P-PSDB; ABR82940.  
 XX  
 PT Producing a transgenic plant having an increased plant resistance, plant  
 PT growth or seed production comprises transforming a plant with a nucleic  
 PT acid molecule having a Proline-rich Extensin-like Receptor Kinase  
 PT activity.  
 XX  
 PS Disclosure; Fig 12; 123pp; English.  
 XX  
 CC The invention relates to producing a transgenic plant having increased  
 CC plant height, number of branches, number of seed pods and/or seed  
 CC production compared to a non-transgenic plant, and/or quicker flowering  
 CC or later senescence compared to a non-transgenic plant. The method  
 CC involves transforming a plant with a vector including a Proline-rich  
 CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic  
 CC acid molecule having PERK activity. The method, as well as the PERK  
 CC nucleic acid molecule and polypeptide, are useful in increasing plant  
 CC resistance to wounding and pathogens and in increasing plant growth and  
 CC seed production. The nucleic acid molecule and polypeptide may also be  
 CC used in producing transgenic plants or transgenic host cells. The present  
 CC sequence represents a PERK1 polypeptide related protein encoding genomic  
 CC DNA from A. thaliana (Accession NO. AAD15491)  
 XX  
 SQ Sequence 2520 BP; 717 A; 552 C; 538 G; 713 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3,58e-36 Length: 2520  
 Score: 1504.00 Matches: 367  
 Percent Similarity: 53.47% Conservative: 80  
 Best Local Similarity: 43.90% Mismatches: 146  
 Query Match: 43.56% Indels: 224  
 DB: 9 Gaps: 22  
 US-10-086-464-2 (1-647) x ACF36551 (1-2520)



```

Db 2139 GCTAGCATTCGTCATTCGGGGCTAAACGTCAAAGATGAG-CCAGGTGAATCAAATTA 2197
QY 543 -----|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2198 TAACTAAAGTCTATTTTGTTCAGAGAATAACAAACAAATGTTGTGGTTTTCAGATAGTA 2257
QY 546 ArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAenGluGlyMetArgProGlyGln 565
Db 2258 AGAGCGTTAGAGGAGAAAGTGTCTTAGATGCTTTTAAACGAAGGTGTGAAGCCAGACAC 2317
QY 566 SerAenValTyrSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAenGlu 585
Db 2318 AGTACGGTTTACGGGTCAATTGGGACGAAGCTCGGATTATAGTCAGACATCTTACAATGCA 2377
QY 586 AspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr-----AsnAla 602
Db 2378 GACATGAAGAAATTCAGACAGATAGCTTTGTGAGCCAAGAAATTCACAGTCAGTGCTGT 2437
QY 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGly 614
Db 2438 GAAGGAACATCTAGTAATGATTCTAGAGATATGGGA 2473
```

RESULT 11

```

ID AAF77097 standard; DNA; 2820 BP.
AC AAF77097;
XX
XX 17-MAY-2001 (first entry)
XX
XX Arabidopsis gene #3.
XX
XX Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
XX
XX Arabidopsis thaliana.
XX
XX WO200114563-A1.
XX
XX 01-MAR-2001.
XX
XX 18-AUG-2000; 2000WO-CA000966.
XX
XX 19-AUG-1999; 99US-0149466P.
XX
XX 13-OCT-1999; 99US-0159122P.
XX.
XX (GORI/) GORING D.
XX (SILV/) SILVA N.
XX
XX Goring D, Silva N;
XX
XX WPI; 2001-244305/25.
XX
XX New proline-rich, extensin-like receptor kinase nucleic acids and
XX polypeptides useful for increasing plant wounding or pathogen resistance,
XX or for producing transgenic plants with increased wounding or pathogen
XX resistance.
XX
XX Example; Fig 13; 91pp; English.
XX
XX The present invention relates to proline-rich extensin-like receptor
XX kinase (PERK). The PERK nucleic acids and polypeptides are useful for
XX increasing the resistance of plants to wounding and pathogens. These are
XX also useful for producing transgenic plants with increased wounding and
XX pathogen resistance compared with a wild type plant, as well as in assays
XX for identifying and developing compounds to inhibit and/or enhance
XX polypeptide function directly
XX
XX Sequence 2820 BP; 819 A; 629 C; 577 G; 795 T; 0 U; 0 Other;
SQ
```

Alignment Scores: 2e-35 Length: 2820  
Pred. No.: 1479.50 Matches: 367  
Score:

```

Percent Similarity: 50.00% Conservative: 99
Best Local Similarity: 39.38% Mismatches: 151
Query Match: 42.85% Indels: 318
DB: 4 Gaps: 28

US-10-086-464-2 (1-647) x AAF77097 (1-2820)

QY 1 MetSerSerAlaProSerProGlyThrGlySerProSerProSerProSerAsnSerThr 20
Db 52 GTGATTTCATCTCTGCCCTGAAACC---TCAATGGGACACCACCGTCAACCGGAACA 108
QY 21 ThrThrThrProProAlaSerAlaProProThrThrProSerSerProProPro 40
Db 109 TCG-----CCGTCTAATGAGTCATCGCGGCCAACACACACCTTCTTCACACCA 159
QY 41 ProSerThrThrProSerProSerSerArgSerThrProSerAlaProPro 60
Db 160 TCATCA-----ATATCTGCTCTCCGCGCAGATATCTCCGCTTCTTTTTCACGCGC 210
QY 61 ProSerProProThr-----ProSerThrProGlySerProProProleu 75
Db 211 CCTGCACCAACGCAAGAAACGTCACCTCTCATATCTCCGCTCATCCGCCCTGTT 270
QY 76 -----ProGln-----ProSerProProAlaPro-----Thr 84
Db 271 GTAGCTAATCCGTCACCGCAGACTCCAGAGAAATCTTCTCCACTGCACCTGAAGGTCA 330
QY 85 ThrProGlySerProProAla---ProValThrPro-----ProThrArgAsn 99
Db 331 ACTCCTGTAAAGCCACCTGCACCAACCAACACCGTCGAACCAATCACCAGGAAGACCA 390
QY 100 ProProPro-----Pro-----Pro-----Pro-----Pro----- 102
Db 391 ACTCCTCTCTCTCTCTGTCCTCAATGATGACCGAAACAGAACCAATGGCGGAAACAAC 450
QY 103 -----SerValProGlyProProSerAsnProSerArg-----GluGlyGly 116
Db 451 AGAGACGGCTCCACCATCATCACCGCTGTCGAGGACAGAACTTCGGGTGACGGTGGC 510
QY 117 SerProArgProProSerProSerProSerProSerProSerAspGlyLeuSerThr 136
Db 511 TCACCTTCACCACTCGGTGATGATGAGCCCTCTCAGAAATAGTGGAGATTCAGACTCATCA 570
QY 137 -----GlyValValValGlyValAlaIleGlyVal 147
Db 571 TCGGGTAATCATCCAAAGCCAAACATTGGATTGATTGAGTCTTGTAGGAGCAGGG 630
QY 148 AlaLeuLeuValIleValThrLeuIleCysLysLeuLeuCysLysLysLysArgArgAsp 167
Db 631 CTTTGTCTTCTACTGTCAGTGTGTATTGTGCATCTGTTGCAACAGGAGAGAGAGAAA 690
QY 168 GluGluAspAla-----TyrTyrValProProProProProProGlyProLysAla 184
Db 691 TCTCCTCAGGTCAACCAACATGCATCTACTACAATAACAATCTTATGGAGGAGCACCCTCA 750
QY 185 Gly-----Pro-----Pro-----Pro-----Pro-----Pro----- 185
Db 751 GGTAATTACAGTTTAGTATATACTCGAAATTAATTTGTAGCCTAATGGTGTGTTGATTAGGT 810
QY 186 -----GlyProTyrGlyGlyGlnGlnGlnGln 194
Db 811 TTCAGAACGATCATAGTCTAATGGTTTCTGCTAGCTCCATATGGCAAGAGGTAGATT- 869
QY 195 TrpArgGlnGlnAenAlaThr-----ProPro 203
Db 870 TATAAGCTAAAGGAGAGTGTTCATAGTGTAGGTAAATGGTGTGTTATTACAAGGAGACCT 929
QY 204 SerAspHisValValThrSerLeuProProProLysAlaProSerProProArgGln 223
Db 930 CAAGATCATGTGGTG---AATATGGCTGGTCAAGAGGTGGGAATTTGGGTGCCACAGCAA 986
QY 224 ProProProProProProProPheMetSerSerGlyGlySerAspTyrSerAsp 243
```

Db 987 CCTGTGCTGGTCTCT-----CACAGTGATGCTTCCAACTTAACCGGT 1028  
Qy 244 ArgProValLeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPhe 263  
Db 1029 CGAACTGCTATACCGTCACTCAA---GCTGCAACTCTTGGTCACAAACCAAGCACCTTC 1085  
Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGln 283  
Db 1086 ACATACGATGAATGCTCCATTGCAACAGAGGTTTCGCTCAGTCAAATTTGCTAGGACAA 1145  
Qy 284 GlyGlyPheGlyTyrValHisGlyGlyValLeuProSerGlyLysGluValAlaValLys 303  
Db 1146 GGAGGATTTGGGTATGTTCAATAAGGAGTCTGCTAGTGGCAAGAGATTTGCGAGTGAAG 1205  
Qy 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeile 323  
Db 1206 AGCTTTAAACTTGAAGTGGACAGGGGAGCGGAGTTTCAAGCAGAGGTTGATATCAAT 1265  
Qy 324 SerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343  
Db 1266 AGCCGTGTCCATCATCGTCATCTCGTTCTCTGTTGGATATTGCATCTCTGGTGGTCAA 1325  
Qy 344 ArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHis----- 361  
Db 1326 AGACTTTTGGTTATGAGTTTATACCTTAACACACTCTTGAATTTTCATCTTCATGGTACA 1385  
Qy 361 ----- 361  
Db 1386 TTCATCTAACAGAAATGTTTCTTGATTAACAAACCTTTAAAGTATGTTTCTCTTTAAT 1445  
Qy 362 -----GlyGluGlyArgProThrMetGluTrpSerThrArgLe 374  
Db 1446 CAGGAACATGATTGAAATTTTCAGGAAGAGGTCGTCGGTTTGGATTGGCCCTACAGAGT 1505  
Qy 374 uLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys----- 391  
Db 1506 GAAGATTGCATTTGGGATCAGCTAGAGGCTTTGCATATTTTGCATGAAGACTGTAAAGAAAT 1565  
Qy 391 ----- 391  
Db 1566 CTTTATCTCACATATTGTCATCAGTTTCTATCTCGCTCTTACATATTTGAAAGATTGT 1625  
Qy 392 -----AsnProLysIleIleHisArgAspIleLysAlaSerAs 404  
Db 1626 ATATGTTACATCAATTATAGTCACTCCCTCGCATTTATCCACAGATATCAAGCTGCAAA 1685  
Qy 404 nIleLeuLeaAspPheLysPheGluAla----- 413  
Db 1686 CATCTCTTCTGATTTTCAGTTTGTAGAC-CAAGGTATGTGTGTATATATATCGACTCTTGTAC 1744  
Qy 413 ----- 413  
Db 1745 TACTTTTACTTTCATGTCCTCTCATTTTGTTCCTCAATCTGTCGATGTGTATCAGT 1804  
Qy 414 -----LysValAlaAspPheGlyLeuAlaLysIleAlaSerAspTh 427  
Db 1805 CTTATTGTGTAATAATATGTCAGGTGGCAGATTTTGGATTGGCTAAGCTATCTCAAGACAA 1864  
Qy 427 rAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440  
Db 1865 CTATACTCATGTCCTCCACTCGCGTCATGGGAACCTTTTGGGTAAAGCAGCTTTGTAAAAATGT 1924  
Qy 440 ----- 440  
Db 1925 CTCAACTCATCCACACTTATTTAGTTTCTTTCACCTGTTTTCATTTTAACTTTCTTGGATTCA 1984  
Qy 441 --TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheS 460  
Db 1985 GATACTTAGTCCAGAGTATGTCATCAAGCGGAAAGTTATCCGACAAATCTGATGTTTCT 2044  
Qy 460 erPheGlyValValLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnV 480  
Db 2045 CATTTGGAGTAATGCTTCTTGAGCTCATAAACCGGAAGACCTCTCTCGATCTAACTCGGA- 2103

Qy 480 alTyrValAspAspSerLeuValAspTrp----- 489  
Db 2104 --GAAATGGAGATAGCTTGGTAGATTGGTAAGTCCGCTCCCGCTCTTTCGGTTTACTTT 2161  
Qy 489 ----- 489  
Db 2162 GTTTAATCCCAAAACACTTTCCAAAGCAAAACAGAAACAAATCTTACTATTGTTGTGTC 2221  
Qy 490 --AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspA 509  
Db 2222 AGCAAGGCTTTGTTGTTGAAAGCAGCTCAAGATGGAGATTACAACCAATTTGCTGATC 2281  
Qy 509 lAlysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaA 529  
Db 2282 CACGCTAGAGCTAAACTCAGTCATCAGAGATGGTTCAAATGGCTTCTTGTGCGAGCTG 2341  
Qy 529 laCysValArgHisSerAlaArgArgProArgMetSerGlnIle----- 544  
Db 2342 CAGCAATCAGACATTCAGCAAGAGACGGCTTAAGATGAGCCAGGT-TCAAAAAACTCATA 2400  
Qy 544 ----- 544  
Db 2401 CCACCTTGTGTTCTATTGTTATATTTTACTACAATTAATCTTGATGATAAATGTGAC 2460  
Qy 545 -----ValArgAlaLeuGluG 550  
Db 2461 ATACTAATGAATCTTGAACATGTGTATGTTAAATGAAGAGATTGTACGACACTAGAG 2520  
Qy 550 lyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrS 570  
Db 2521 GAGATATGTCAATGGATGATCTAAGTGAGGGAACAACACAGGACAAAGCAGCTACTTGA 2580  
Qy 570 erSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysP 590  
Db 2581 GCCCGGAGCGTGCAGCTCAGAGTATGACGCAAGCTCGTACCGGCAGACATGAAAAAGT 2640  
Qy 590 heArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGlyTyrSerAsnProT 610  
Db 2641 TCAAGAAACTGGGCTTAGAGATAAGAAATATCAAGCAGT---GAATATGGTGA---A 2694  
Qy 610 hrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnTyrThrArgGluM 630  
Db 2695 CAAGTGATGATGCTTAAACCTTCTGCTTCAAGTAGTGAA-----GAAA 2739  
Qy 630 etGluMetGlyLysIleLysArgThrGlyGln 640  
Db 2740 TGAATAGAGGCTCAATGAAACGCAATCCTCAG 2771  
RESULT 12  
ACF36552  
ID ACF36552 standard; DNA; 2820 BP.  
XX AC ACF36552;  
XX AC  
DT 18-DEC-2003 (first entry)  
XX DE  
XX Arabidopsis PERK1 receptor related protein encoding DNA.  
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;  
KW wound; pathogen resistance; plant growth; seed production; gene; ds.  
XX OS Arabidopsis thaliana.  
XX PN WO2003072763-A1.  
XX PD  
XX 04-SEP-2003.  
XX PF 28-FEB-2003; 2003WO-CA000274.  
XX PR 28-FEB-2002; 2002CA-02373903.  
XX PR 28-FEB-2002; 2002US-00086464.  
XX XX



```
Qy 391 ----- 391
Db 1566 CTTTATCTCACATATTTTCATCAGTTTCTATCTCGCTCTCTACATAATTTTCAAAGATTGT 1625
Qy 392 -----
Db 1626 ATATGTTACATCAATATAGTCACCTCGCATATTCACAGAGATATCAAAAGCTGCAAA 1685
Qy 404 nileuileaspPheLysPheGluAla----- 413
Db 1686 CATCTCTTGATTTTCAGTTTGTAGAC--CAAGGTATGTGTATATATCGACTCTTGTAC 1744
Qy 413 ----- 413
Db 1745 TACTTTTACTTTTCATTTGCTCTCTATTTTGTTCCTCAATCTGTGTGATGTGATCAGT 1804
Qy 414 -----LysValAlaAspPheGlyLeuAlaLysLeuAlaLysLeuAlaSerAspTh 427
Db 1805 CTTATTTGTTAAATATATGCGAGGTGGCAGATTTTGGATTGGCTAAGCTATCTCAAGACAA 1864
Qy 427 xAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440
Db 1865 CTATACTCATGTCCTCCACTCGCTCATGGGAACCTTTGGGTAAAGCAGCTTTGTAATATGT 1924
Qy 440 ----- 440
Db 1925 CTCAACTCATCCACACTTATTAGTTTCTTCACTTGTGTTTTTAACATTTTCTTGGATTCA 1984
Qy 441 --TyrLeuAlaProGluTyrAlaLaserGlyLysLeuThrGluLysSerAspValPheS 460
Db 1985 GATACTTAGCTCCAGAGTATCATCAAGCGGAAAGTTATCCGACAAATCTGATGTTTCT 2044
Qy 460 erPheGlyValValLeuLeuGluLeuThrGlyArgProValAspAlaAsnAsnV 480
Db 2045 CATTTGGAGTAATGCTTCTGAGCTCATACCGGAAGACCTCTCTGGAATCACTGGA- 2103
Qy 480 alTyrValAspAspSerLeuValAspTrp----- 489
Db 2104 --GAAATGGAAGATAGCTTGTAGATTGGTAAAGTTCGGTCCCGCCTCTTCGGTTTACTT 2161
Qy 489 ----- 489
Db 2162 GTTTAATCCCAAAACACTTTCCAAAGCAAAACAGAAACAAATCTTACTATTTGTTGTC 2221
Qy 490 --AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspA 509
Db 2222 AGCAAGCCCTTTGTGTTGAAGAGCGCTCAAGATGGAGATTAACCAATTTGGCTGATC 2281
Qy 509 laLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaA 529
Db 2282 CACGCTAGACCTAAACTACAGTCATCAAGAGATGGTTCAATGGCTTCTTGTGCAGCTG 2341
Qy 529 laCysValArgHisSerAlaArgArgProArgMetSerGlnle----- 544
Db 2342 CAGCAATCAGACATTCAGCAAGAAGACGGCTTAAGATGAGCAGCTG-TCAAAAACCTCATA 2400
Qy 544 ----- 544
Db 2401 CCACCTGTTGTTCTATTTGTTATATTTTACTCAAAATAATCTTGATGATAAATGTGAC 2460
Qy 545 -----ValArgAlaLeuGluG 550
Db 2461 ATACTAATGAATCTTGAAACATGTGTATGGTAAATGAAAGATTGTACGAGCACTAGAAG 2520
Qy 550 lyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrS 570
Db 2521 GAGATATGTCAATGATGATCTAAGTGAAGGAACAAGACCAAGGACAAAGCAGCTACTTGA 2580
Qy 570 erSerTyrGlyLysSerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysP 590
Db 2581 GCCCGGAGCGGTGAGCTCAGAGTATGACGAAGCTGTGTACCGCGACATGAAAGAT 2640
Qy 590 heArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProT 610
```

```
Db 2641 TCAGAAACTGGCGTTAGAGAATAAAGATATCAAGCAGT--GAATATGGTGA---A 2694
Qy 610 hrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThrArgGluM 630
Db 2695 CAAGTGAATATGGCTTAAACCCCTTCTCTCAAGTAGTGAA-----GAAA 2739
Qy 630 etGluMetGlyLysIleLysArgThrGlyGln 640
Db 2740 TGAATAGAGGCTCAATGAAGCAATCCTCAG 2771
RESULT 13
AAF77095
ID AAF77095 standard; DNA; 2880 BP.
XX
AC AAF77095;
XX
DT 17-MAY-2001 (first entry)
XX
DE Arabidopsis gene #1.
XX
KW Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.
XX
OS Arabidopsis thaliana.
XX
FN W0200114563-Al.
XX
PD 01-MAR-2001.
XX
PF 18-AUG-2000; 2000WO-CA000966.
XX
PR 19-AUG-1999; 99US-0149466P.
XX
PR 13-OCT-1999; 99US-0159122P.
XX
(GORI/) GORING D.
PA (SILV/) SILVA N.
XX
Goring D, Silva N;
XX
WPI; 2001-244305/25.
XX
New proline-rich, extensin-like receptor kinase nucleic acids and
polypeptides useful for increasing plant wounding or pathogen resistance,
or for producing transgenic plants with increased wounding or pathogen
resistance.
XX
Example; Fig 11; 91pp; English.
XX
The present invention relates to proline-rich extensin-like receptor
kinase (PERK). The PERK nucleic acids and polypeptides are useful for
increasing the resistance of plants to wounding and pathogens. These are
also useful for producing transgenic plants with increased wounding and
pathogen resistance compared with a wild type plant, as well as in assays
for identifying and developing compounds to inhibit and/or enhance
polypeptide function directly
XX
SQ Sequence 2880 BP; 814 A; 693 C; 569 G; 804 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.5e-27 Length: 2880
Score: 1209.50 Matches: 325
Percent Similarity: 43.85% Conservative: 92
Best Local Similarity: 34.17% Mismatches: 160
Query Match: 35.03% Indels: 376
DB: 4 Gaps: 26
US-10-086-464-2 (1-647) x AAF77095 (1-2880)
Qy 2 SerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThrThr 21
Db 98 AACTCAGCTTCTCCACCTGTTGATTCCTCTCTCTAGTCCACGCTGATTCATCA 157
Qy 22 Thr-----ThrProProAlaSer----- 28
```

Db 158 ACACGCCGCTGTGAGAACCATCCACTCTCTCCAGATTACAGCTTCCTCCTTACCT 217  
|||  
Qy 29 -----AlaProProThrThrProSerSerProPro 40  
:::|  
Db 218 TCGATTCTTCCTCCGCTAACAGATTCTCCACCTCCACCTTCGATTTCTTCCACCGTT 277  
:::|  
Qy 41 ProSerThrIleProThrSerProProProSerSerArgSerThrProSerAlaPro--- 59  
|||:::|  
Db 278 GATTCAACC---CCTTCTCCGCGCCACGAGCTCAAAACGAATCTCTTCTCTCCAGAA 334  
|||:::|  
Qy 60 -----ProProProProProThrProProGlySerProProPro--- 74  
|||:::|  
Db 335 GATTCGGAACACCCACCTGCTCCACCAATGAATCCAAATGACACCAACCTCTCTCCGCT 394  
|||:::|  
Qy 75 -----LeuProGlnProSerProProAla-----ProThrThrPro 86  
|||:::|  
Db 395 CAAGATCTTCAATCGCTCTCCATCGTCCGCGTCCGGAATGTAGGACCCACAAACCG 454  
|||:::|  
Qy 87 GlySerProPro-----AlaProValThrProProThrArgAsnPro-----Pro 101  
|||:::|  
Db 455 GAATCACCACCGTTACAAATCTCTCCAGCTCCACCAGCATCAGATCTTACAAATTCACCG 514  
|||:::|  
Qy 102 ProSerValProGlyProProSerAsnPro-----SerArgGluGlyGlySerPro 118  
|||:::|  
Db 515 CCAGCTTCCACCTTACAGCTTACCAATCTCTCCCAATACAAACCATCAGGACCGCACT 574  
|||:::|  
Qy 119 ArgProProSerProPro-----SerProProSerPro----- 129  
|||:::|  
Db 575 TCTCTCCGCTAATCCCAAGCTCCGCGGAGCCATTCGCCACAGTACCACCCCAAACT 634  
|||:::|  
Qy 129 ----- 129  
|||  
Db 635 CTTCTAGTGAGCTGTGTCTCCATCTCTCACATCCCTAGTAAAGAACTCCTACT 694  
|||  
Qy 130 -----SerSerAspGlyLeuSerThrGly-----ValVal 139  
|||  
Db 695 CCMAACCAAGCAATGGAGATGGCGGTGGCGGTGTGGCGGTATCAAGGGAAGACTATG 754  
|||  
Qy 140 ValGlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeu 159  
|||:::|  
Db 755 GTTGATGTGCTGATCGCGTTTCGCAATCATGGCGTTATAGGCGTTGTGTCTTAGTG 814  
|||  
Qy 160 CysLysLysLysArgArgAspGluGluAspAlaTyr-----TyrValPro 175  
|||:::|  
Db 815 AGAAGAAAGAAAGAGA-----AACATTGATGCTATAATCACTCACAGTACTTGCCA 868  
|||  
Qy 176 ProPro----- 177  
|||  
Db 869 CATCCAAATTTCTGTTAAATCAGGTTTAAATAATCTCACCTTTATCTCTCTGATCAT 928  
|||  
Qy 177 ----- 177  
|||  
Db 929 CTTCTATGTGTTGAATCATCTCTGACTATCTTTGCTTTTGTAGTAGATGATTTCTTA 988  
|||  
Qy 178 -----ProProProGlyProLysAlaGlyProTyrGlyGlyGlnGlnGln 194  
|||:::|  
Db 989 TACGCTCAAGATCCAGGTAAGGATPACTCCTCTGCTCTCTTAATGGTTCAATGTATAACAAT 1048  
|||  
Qy 195 TrpArgGlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuProProPro 214  
|||:::|  
Db 1049 TCACAGCAACAACATCTCTATGGGAACAGTTATGTTAGTACAGCTGGT----- 1096  
|||  
Qy 215 ProLysAlaProSerProProArgGlnProProProProProProPheMetSer 234  
|||  
Db 1097 -----GGTGGTTATCTCTCATCA-----ATGCAA 1123  
|||  
Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProProSerProGlyLeu 254  
|||  
Db 1124 TCAAGTGGACACCTGC-----TCTGCT 1147  
|||  
Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274  
|||:::|

Db 1148 ATACTCGGAAGTGCCAGACTCAATTCAGTTACGAGAGCTTGCTGAGATAACACAAGGC 1207  
|||  
Qy 275 PheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeu 294  
|||:::|  
Db 1208 TTTCTCGCAAAAACATTTCTTGGAAGCGGATTTGGATGTGTCTATAAAGGTACATTG 1267  
|||  
Qy 295 ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArg 314  
|||:::|  
Db 1268 CAGATGTTAAAGTTGTTGCGTTAAGAGCTTAAAGCTGGAAGTGGACAAGGTGACCGT 1327  
|||  
Qy 315 GluPheGlnAlaGluValGluIleSerArgValHisHisArgHisLeuValSerLeu 334  
|||:::|  
Db 1328 GAATTCAGACAGAGTTGAGATCATCAGCGCGTTTCATCATCGCATTTGCTCTCTG 1387  
|||  
Qy 335 ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn 354  
|||:::|  
Db 1388 GTTGGTTACTGCAATTCAGACCAGCATAGATTGCTTATCATATGATGATGTTTCTAATCAA 1447  
|||  
Qy 355 AsnLeuGluLeuHisLeuHisGlyGlu----- 363  
|||  
Db 1448 ACCTGGAGCATCAATTTGCATGGTGAGTGACTGTTACCATTTCGTTATAGATAAGACT 1507  
|||  
Qy 364 -----GlyArgProTh 367  
|||  
Db 1508 TTTTGTAGCTTTACGTTTAGACTGACTCGCTTTACGCTTTAGGAAAGGTTTGCAGT 1567  
|||  
Qy 367 rMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLe 387  
|||:::|  
Db 1568 TTTAGATGGTCTAAGAGAGTCCGGATCGCTATAGGATCAGCCAAGGGTTGGCATATCT 1627  
|||  
Qy 387 uHisGluAspCysAsn----- 392  
|||  
Db 1628 TCACAAAGCTGTAA--GTAATGCCTTCACATTTTCTTAGTTGTGTGCTTTGGTTATGCAC 1686  
|||  
Qy 393 -----ProLy 394  
|||  
Db 1687 TTCATAGTTTAAACAGAAAGCCAAAATCATATCTCTGTTTATTTTACAGGTCATCCGA 1746  
|||  
Qy 394 sIleIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLy 414  
|||:::|  
Db 1747 AATCATTCACAGATATATAAGTCAGCAATATTTCTTCTAGATGATGAATATGAAGCTCA 1806  
|||  
Qy 414 s----- 414  
|||  
Db 1807 -GGCAATAATGAATAATCTCTCTTTTCGTTAAATCTATCTTATGACTGTAAAGTTTAGTTA 1865  
|||  
Qy 415 -----ValAlaAspPheGlyLeuAlaLysIleAl 424  
|||  
Db 1866 ATGAGACTTGTCTGTTTTTTTGGATGTTTAGTTGCTGATTTTGGACTTGTAGACTCAA 1925  
|||  
Qy 424 aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly----- 440  
|||  
Db 1926 TGATAACAACAACACTCATGTTTCACTCGGGTTATGGGAACCTTCGGGTAAAGCAACAT 1985  
|||  
Qy 440 ----- 440  
|||  
Db 1986 TCATCAAACTCTACTCCAAAACCTGGACCTTATTGATCCAATGCTGATGAAAAGTTTG 2045  
|||  
Qy 440 ----- 440  
|||  
Db 2046 TTATATATGGCTTGAGGCAACAAATTTGGATCAAACTGAAATCTTTTATGATCGTATGGCT 2105  
|||  
Qy 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuT 453  
|||  
Db 2106 GCATGACATGTTTGTGTTAAGGTACCTAGCGCGGAATATGATCAAGTGGAAATTTGA 2165  
|||  
Qy 453 hrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 473  
|||:::|  
Db 2166 CTGATAGATCGATGTTATCTCATTCGGGGTGTCTCTTAGAGCTTGTAACTGGACGGA 2225  
|||  
Qy 473 rgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp----- 489  
|||:::|  
Db 2226 AACCAAGTTGACAGACTCAGCCTCTAGGAGAAGAGAGTTTGGTTGAATGGGTAAAGATCC 2285  
|||



```
QY 130 -----SerSerAspGlyLeuSerThrGly-----ValVal 139
Db 695 CCAACCAAGCAATGGAGATGGCGTGGCGGTATCAAGGGAAGACTATG 754
QY 140 ValGlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeu 159
Db 755 GTTGTATGGCTGTAGCGGTTTCGAATCATATGGCGCTTATAGCGGTGTGTCTTAGTG 814
QY 160 CysIlySlyAsArgArgAspGluGluAspAlaTyr-----TyrValPro 175
Db 815 AGAAGAGAGAAAAGAGA-----AACATTGATAGCTATATCACTCACAGTACTTGCCA 868
QY 176 ProPro----- 177
Db 869 CATCCCAATTCTCTGTATAATCAGGTTTAAAAATCTCACCTTTATCTCTCTCTCATCAT 928
QY 177 ----- 177
Db 929 CTTCTATGTGCTTGAATCATCTCTGACTATCTTTTGTGTTTGTATGTATGGATTTCTTA 988
QY 178 -----ProProProGlyProlysAlaGlyGlyProTyrGlyGlyGlnGlnGln 194
Db 989 TACGGTCAAGATCCAGGTAAGGATACCTCTCTGCTCTAATGGTTCAATGTATACAAT 1048
QY 195 TrpArgGlnAsnAlaThrProProSerAspHisValValThrSerLeuProPro 214
Db 1049 TCACAGCAACAACATCTCTATGGAAACAGTTATGTGTACAGCTGGT----- 1096
QY 215 ProlysAlaProSerProArgGlnProProProProValLeuProProPheMetSer 234
Db 1097 -----GGTGGTATCTCTCATCATCAA-----ATGCCAA 1123
QY 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProProSerProGlyLeu 254
Db 1124 TCAAGTGCACACCTGAC-----TCTGCT 1147
QY 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
Db 1148 ATATCGGAAGTGGCCAGCATTTTCAAGTACGAAGAGCTTGCTGAGATAACACAGGC 1207
QY 275 PheSerGluAlaAsnLeuGlyGlnGlyPheGlyTyrValHisGlyValLeu 294
Db 1208 TTTGCTCGCAAAACATTTCTGGAGAGCGGATTTGGATGTCTATAAAGGTACATTTG 1267
QY 295 ProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArg 314
Db 1268 CAGGATGGTAAAGTTGTTCCGGTTAAGCAGCTTAAAGCTGGAGTGGACAGGTGACCGT 1327
QY 315 GluPheGlnAlaGluValGluIleIleSerArgValHisArgHisLeuValSerLeu 334
Db 1328 GAATTCAAAGCAGAGGTTGAGATCATCAGCGCGTTTCATCATCGCCATTTGCTCTCTG 1387
QY 335 ValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsn 354
Db 1388 GTTGGTTACTGCATTTTCAGACAGCATAGATTGCTTATCTATGATGATGTTTCTTAATCAA 1447
QY 355 AsnLeuGluLeuHisGlyGlu----- 363
Db 1448 ACCTTGGAGCATCATTTGATGGTGAGTGACTTTGTACCATTTTCGTTATAGATAAGACT 1507
QY 364 -----GlyArgProTh 367
Db 1508 TTTTATAGCTTTTACGTGTTAGACTGACTCGCTTTTACGCTTTAGGAAAGGTTTGCAGT 1567
QY 367 rMetGlnTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLe 387
Db 1568 TTTAGATGGTCTTAAGAGAGTCCGATCGCTATAGGATCAGCAAGGGTTGGCATATCT 1627
QY 387 uHisGluAspCysAsn----- 392
Db 1628 TCACGAGACATGTA -GTAATGCCCTTCACATTTTCTTAGTTGTGTGCTTTGTTATGCAC 1686
```

```
QY 393 -----ProLy 394
Db 1687 TTCATAGTTTAAACAGAACGCAAAATCATATCTTGTGTTTATTTTACAGGTCTATCCGAA 1746
QY 394 sIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLy 414
Db 1747 ANTCAATTCACAGATATATAAGTCAGCAATATTTCTTCTAGATGATGAATATGAAGCTCA 1806
QY 414 s----- 414
Db 1807 -GGCAATAATGAAATCCTCTTTTCGTTAAATCTATCTTATGACTGTAAAAGTTTAGTTA 1865
QY 415 -----ValAlaAspPheGlyLeuAlaLysIleAl 424
Db 1866 ATGAGACTGTGTTCTGTTTTTTTGGATGTTAGTTGCTGATTTTGGACTTGTAGACTCAA 1925
QY 424 aserAspThrAsnThrHisValSerThrArgValMetGlyThrPhe-Gly 440
Db 1926 TGATACAACACAAACTCATGTTTCACTCGGGTTATGGGAACCTTCGGGTAAAGCAACAT 1985
QY 440 ----- 440
Db 1986 TCATCACAACACTCTACTCCAAACTCGGACCTTATTGATCCAATGCTGATGAAAAAGTTTG 2045
QY 440 ----- 440
Db 2046 TTATATATGGCTTGAGGCAACAAATGGATCAAACTGAAATCTTTTATTTGATCGTATGGCT 2105
QY 441 -----TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 453
Db 2106 GCATGACATGTTTGTGTTAAGTACCTAGCGCCGGAATATGCATCAAGTCGAAAAATTGA 2165
QY 453 hrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgA 473
Db 2166 CTGATAGATCGATGTTCTTCAATTCGGGGTTGTTCTCTTAGAGCTTGTAACTGGACGGA 2225
QY 473 rgProValAspAlaAsnAsnValTyrValAspAspSerLeuValAspTrp----- 489
Db 2226 AACCAGTTGACAGACATCAGCCTCTAGGAGAAGAGAGTTTGGTGAATGGGTAAAGAATCC 2285
QY 489 ----- 489
Db 2286 AACTTTCAACATTTCTCAATAATAGTAGATTGGCCCTAGTATATCTATATATAGTACTTA 2345
QY 490 -----AlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlu 504
Db 2346 TAAATGAACATCACAGGCGCGCCGCTCTCTCAAAGCCATTGAGACCGGAGATTTAAGC 2405
QY 505 GlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMetAlaArgMetVal 524
Db 2406 GAATCTGATTGATACACGCTTTGAAAGCGTTTATGTGGAGCATGAAGTCTTCAGAAATGATC 2465
QY 525 AlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGln--- 543
Db 2466 GAGACAGCGCTGCATGTGTGTAGACATCTCGTCCAAAACGCCACGATGTTTCAGGTA 2525
QY 543 ----- 543
Db 2526 ATTCTGACTAACCAAAAGTCCAAAAGTCCCATATATAGTAACAAGTGATTTCTCACATCT 2585
QY 544 -----IleValArgAlaLeuGluGlyAsnValSerLeu 554
Db 2586 GAAAACTTATCTACTCTTCGAAATAAGGTTGTGAGACATTTGGACTGCGCGGAGACTCG 2645
QY 555 SerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGly 574
Db 2646 GGATATATGACACCGGATCAAAATTGGCAA----- 2678
QY 575 SerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAla 594
Db 2679 AGCACACTTATGACTCAGGGCAATACAATGAAGACATTTATGAAATTCAGGAAATGGCG 2738
QY 595 LeuGlyThrGlnGluTyrAsnAlaThrGlyLutyrSerAsnProThrSerAspTyrGly 614
```

Db 2739 TTTGGTGTGATAACAGCGTAGAGTCAGGATTGTACAGT-----CGA 2780  
||| :||| |||||

Qy 615 LeuTyrProSerGlySerSerSerGlu 623  
||| :||| |||||

Db 2781 AACTACTCTGCCAAAGCTCTTCAGAT 2807  
||| :||| |||||

## RESULT 15

ID ACF36553 standard; DNA; 3060 BP.

XX ACF36553;

AC 18-DEC-2003 (first entry)

DT Arabidopsis PERK1 receptor related protein encoding DNA.

XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;

XX wound; pathogen resistance; plant growth; seed production; gene; ds.

XX Arabidopsis thaliana.

XX WO2003072763-A1.

XX 04-SEP-2003; 2003WO-CA000274.

XX 28-FEB-2003; 2002CA-02373903.

PR 28-FEB-2002; 2002US-00086464.

XX (GORI/) GORING D.

PA (SILV/) SILVA N.

PA (HAFF/) HAFFANI Y Z.

XX Goring D, Silva N, Haffani YZ;

XX WPI; 2003-712727/67.

DR P-PSDB; ABR82942.

XX Producing a transgenic plant having an increased plant resistance, plant growth or seed production comprises transforming a plant with a nucleic acid molecule having a Proline-rich Extensin-like Receptor Kinase activity.

XX Disclosure; Fig 14; 123pp; English.

XX The invention relates to producing a transgenic plant having increased plant height, number of branches, number of seed pods and/or seed production compared to a non-transgenic plant, and/or quicker flowering or later senescence compared to a non-transgenic plant. The method involves transforming a plant with a vector including a proline-rich Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic acid molecule having PERK activity. The method, as well as the PERK nucleic acid molecule and polypeptide, are useful in increasing plant resistance to wounding and pathogens and in increasing plant growth and seed production. The nucleic acid molecule and polypeptide may also be used in producing transgenic plants or transgenic host cells. The present sequence represents a PERK1 polypeptide related protein encoding genomic DNA from A. thaliana (Accession NO. CAA18590)

SQ Sequence 3060 BP; 797 A; 766 C; 592 G; 905 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	8.1e-27	Length:	3060
Score:	1185.00	Matches:	342
Percent Similarity:	39.28%	Conservative:	72
Best Local Similarity:	32.45%	Mismatches:	185
Query Match:	34.32%	Indels:	456
DB:	9	Gaps:	31

US-10-086-464-2 (1-647) x ACF36553 (1-3060)

Qy	2	SerSerAlaProSerProGlyThrGlySerProPro	-----	13
Db	43	TCCTTCCTCGCGGTTCGGCAACT---TCCCGGCTGCATGTACCTACCTCCGGCAGAT	99	
Qy	14	-----SerProSerAsnSerThrThrThrThrProProProAla	27	
Db	100	TCCGTACCTGACACGTCATCACCCTCCAGCTCCTCTTTGTCTCTCTTCCCGCCACCAT	159	
Qy	28	SerAlaProProThrThrProSerProPro	-----	39
Db	160	AGCTCTCTCCGCG---TTGCGCTTACACCGGCTCTCTCCGGTCCACCGCTTCCCGCA	216	
Qy	40	-----ProProSerThrProThrProSerProProProSerSer	52	
Db	217	CCGCTCTTCGCGTTGAATCCCGCTCTCTCTATAGAAATCACCACCGCTCTCTCTA	276	
Qy	53	ArgSerThrProSerAlaProPro-----ProSerProThrPro-----Ser	67	
Db	277	CTGGAATCAGCTCTCTCTCTCTCGTTGGAATCTCATCTCCACCGTCTCTCTCAGCTCA	336	
Qy	68	ThrProGlySerProProLeuPro-----GlnProSerProProAlaPro	83	
Db	337	GCTCCTTCGCGTTACCGGCTTACCTTCTCCCGCCCAACCTTCTCCGCGC-----	390	
Qy	84	ThrThrProGlySerProProAlaProValThrProPro-----ThrArgAsnProPro	101	
Db	391	-----CCTTCTTACCTCTCCGAGACAGATTCCGCGCGGGAATACGATTCTCCACCA	444	
Qy	102	Pro-----SerValProGlyProPro	108	
Db	445	CCTCGTTTCACTTCTCCCGAATCAACCCCGCGGTGAACACAGCTTCTCTCCACCGCA	504	
Qy	109	SerAsnProSerArgGluGlyGlySerProArgPro-----	120	
Db	505	TCCTCTCTCGCGCGCTAGTGGC---CCTAAGCTTCTGTTCTCTCTCCCATCAATTCT	561	
Qy	121	---ProSerSerProPro-----	126	
Db	562	TCTCCACCAATCTTCTCCGAACTCCGCTCCAGAACTTCTCTCCACCTAAA	621	
Qy	127	-----ProSerProSerSer-----	131	
Db	622	CCACCGCTCTCAACGACGCCAATTTCTCTCTCTCACTCCCGCGCTAAGAAGTCCCT	681	
Qy	132	-----AspGlyLeu-----	134	
Db	682	GCAGCAGTAATCTTCTTCTTTGGCCAGCGGGCCAATTCACGGATGGACCGTAGCA	741	
Qy	134	-----	134	
Db	742	CCTCTATTGGGCTGTTATAGAGCCCAAGACGAGTCCAGCCGAATCAATATCTCCGGA	801	
Qy	135	-----Ser	135	
Db	802	ACGCCACGACCTGGTTCCGAAGAGTCTACTGTAAACGACGCTGTATCACCAGTATCC	861	
Qy	136	ThrGlyValValValGlyIleAlaIleGlyValAlaLeuValIleValThrLeu	155	
Db	862	GCCGGATTTCTTTTGGCGGTGAATCGTTGGAGCTTCTCTCTAAATTTCTGTAGGCTT	921	
Qy	156	IleCysLeuLeuCysLysLysAspArgArgAspGluGluAsp-AlaTyrTyrValPr	175	
Db	922	CTCTTTGTCTTCTACAGAGCTACCAGAAATAGAAATAACACACAGCAGCTCTGCTCATCAT	981	
Qy	175	oProProProPro-----	180	
Db	982	CAATCCAAAACCTCCCTCAAAAGGTATAAATTTTGAGATCAATTTGTTTTCAGACTGTCACT	1041	
Qy	181	-----GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnTpar	196	
Db	1042	TAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG	1101	
Qy	196	GlnGlnAsnAlaThrProProSerAspHisValValThrSerLeuProProProProLy	216	

			::	:::	1102	GCGCGGTAATGCCTGGTACGAACCAGGCACATGTTATTACACA---ATGCCACCAACCAATCCA	1158
Db							
Qy	eAlaProSerProProArgGlnProProProProProPheMetSerSerSe	236					
Db	TGCTAAA-----	1159					
Qy	rGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeuValle	236					
Db	TGGAGGTGTGATACGAAGCAGAAC-----AATTCtGTTCGCAAAAAACAT	1221					
Qy	uGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThraSngLyPheSe	276	:	:::			
Db	TTCAATGCCATCTCGAAATGTTCTCCACGAGAACCTTTCAAAAGCAAACCTGGTGCGATTTTC	1281					
Qy	rGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProse	296					
Db	AGAGGAGNACTTTTGGAGAGCGGGTTCCGATATGTTCCANAGGAGTGTGTGA AAAA	1341					
Qy	rGlyLysGluValalaVallysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPh	316					
Db	CGGCACAGAAGTTGCGGTGAAGCAGCTGCAAGATTGGGAGCTATCAAGGGGAAAAGAGAA TT	1401					
Qy	eGlnAlaGluValGluIleSerArgValHisHsArgHisLeuValSerLeuValGl	336					
Db	CCAAGCTCGAGGTTCACACAATCATAGTAGGGTTTCATCATTAAGCACCTCGTTTCAT TGGTGG	1461					
Qy	YTyrcysileAlaGlyAlalaysArgLeuLeuValTyrGluPheValProAsnAsnLe	356					
Db	TTATTGGTGTAAATGGAGATAAAGACTCTTGGTTTACGAGTTTGTTCCTTAAGATACCTT	1521					
Qy	uGluLeuHisLeuHisGly-----	362					
Db	GGAGTTCCACTTCGATCGTAAAAATAGATATATGATTTTCATCTTTTTTGATTTTGTCTTTA	1581					
Qy	-----GluGlyArg---ProThrMetGl	369					
Db	GTTCATTATGTTGAGTATTGTGAGATPATGTGTTGTAGAGAACAGAGGAAGCGTGTGGA	1641					
Qy	uTPserThrArgLeuLysileAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGl	389					
Db	ATGGRAATCAGGCTCAGGATTGCTGTAGGAGCAGCAAAAGGATTAGCTTATCTTCATGA	1701					
Qy	uAppCys-----	391					
Db	GGATTGTGAGTTGTTCTCTTCATAATFGGAATGACAAATGCCCTTTTTTGGGTTTTTCGAC	1761					
Qy	-----AsnProLysileileHisArgAspil	400					
Db	ACTGATATTGATTTCTGGTGTGTATGCAGGCAGTCCAACTATATATTCACCGTGATAT	1821					
Qy	eLysAlaSerAsnileLeuleAspPheLysPheGluAlaLys-----	414					
Db	CMAAGCAGCTAATATCTCTCTAGATTCCAAAATTTGAGCCAAAGGTGATCTGCTCTTTAAT	1881					
Qy	-----	414					
Db	CTTATCAAAGTTGGTTTTTTAGAACAGAGTTTGCCATTTTTTCGGTTTATAGTCACACCAT	1941					
Qy	-----ValAlaAs	417					
Db	TTTGTCTATTAAACAGATTACATGTAGCTTAGAGTTCTTGTCTCTTCAGGTCTCTGA	2001					
Qy	pPheGlyLeuAlaLysileAlaSerAspThrAsn-----ThrHisValSerThrAr	434					
Db	CTTTGGACTAGCCAAAGTTTTTCTCAGACACCAAAATTCATTCATTCATCATATCTCTACTCG	2061					
Qy	gValMerGlyThrPheGly-----	440					
Db	AGTGGTAGGAACCTTCGGGTAAACACCATCCATCCATGCTTTATATATGTTGTGTCAATGT	2121					
Qy	-----	440					

D	b	2122	GTTTAAATAATTAATTACGGTTTCAGGTTTCAAGTTTCAAGTTTCAAGTTTCTTCCTT	2181
Q	y	440	-----	440
D	b	2182	GTTGTATCGTAATCCAGATCAAAGAATTTATTGATTACTAAATGCCTTGTCACCGTCT	2241
Q	y	441	-----TyrLeualaProGluTy rAlaAAserGlyLysLeuth	453
D	b	2242	ATTTGGTATGCAATTTAAACACAGATACATGGCTCCAGAATAACGCCGTCCAGTGGTAAAGTAAC	2301
Q	y	453	rGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeulethrGlyArgAr	473
D	b	2302	TGATAAATCAGATGATATTCCTTTGGGTCGTCCTCTTAGAACTCATCTACGAGCTCC	2361
Q	y	473	gProValasPaAlaasnAsnValTy rValAspAspSerLeuValAspTrp	489
D	b	2362	ATCAATTTTCGCCCAAGATTCTTCCACAACCAGAGTTTTAGTAGACTGGGTAAGTCAAAG	2421
Q	y	489	-----	489
D	b	2422	TACATGATGATGATGATACCATTAGGTTTTCTTTTCCCAGTAGTTATTATTAGATGA	2481
Q	y	489	-----	489
D	b	2482	ATGAACCTCAGTATAACTCGCAAGCTTGAAACTTCATTAGACTGTAANTTTTGATTATTC	2541
Q	y	490	-----AlaArgProLeuLeuAsnArgAlasergluGlnGlyAspPheGluGlyLeuAla	508
D	b	2542	CTCAGGCGAGGCCATTGCTTTACGAAGCAATCTCTGAAGAAGTCTTTTGACTTTCTTGTAG	2601
Q	y	508	sPalalysMetAnnAsnGlyTy rAspArgGluGluMetAlaArgMetValAlaCyAla	528
D	b	2602	ACTCAAGTTGGAGAAGAAATACGATCAACTCAGATGGCAACATGGCTGTGTGCTG	2661
Q	y	528	laLaCyValArgHisserAlaArgArgProArgMetSerGlnile-	544
D	b	2662	CTGCTGCATACGCCAATCAGCTTGGCTTCGGCTAGATGAGCCAGGT-CTGAGATTTA	2720
Q	y	544	-----	544
D	b	2721	GTTTAAATACATGTATTTCGTCCATATATCGAAAAAGAACAGCCTAATCCATGAATACAT	2780
Q	y	545	-----ValargAlaleuGluGlyAsnValserleus	555
D	b	2781	TTATATCTTGAAAAAACCCTTGAATAGTAGTACGTGCTCTTGAAGGCGAGGTGGCCCTGA	2840
Q	y	555	eAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTy rSerSerTy rGlyGlys	575
D	b	2841	GAAGGTCGAAGAG-----ACTGGGAATAGCGTACCTATAGCTCTTCTGAAAAACC	2891
Q	y	575	eThrAspTy rAspSerSerglnty rAsnGluAspMetLysLysPheArgLysMetAlal	595
D	b	2892	CGAATGCATC---ACACCAACCGTATGGAAACAAAATAAGAGAGATTTC-----	2935
Q	y	595	euGlyThrGlnGluTy rAsnAlathrGlyGluTy rSerAsnProThrSerAspTy rGlyL	615
D	b	2936	-----GACACAGGTTCAAGCGATGGTTACACTTCAGATATATGGAG	2975
Q	y	615	euty rProSerGlySerSerglnty rGlnThrThrArgGluMetGluMetGlyLysI	635
D	b	2976	TTAACCTTCTCAGTCGAGCAGTGAACATCAACAGAGGTGAATACTTAGTTCACAGGTTCAA	3035
Q	y	635	leiysArgThrGlyGlnGlyTy rSeroGlyProSerLeu	647
D	b	3036	TA-----GGCAAGTTTCCACACAAATTA	3058

Search completed: May 12, 2004, 22:54:05  
Job time : 666 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2004, 22:25:16 ; Search time 3599 Seconds  
(without alignments)  
5368.389 Million cell updates/sec

Title: US-10-086-464-2  
Perfect score: 3453  
Sequence: 1 MSSAPSGTGGPPSPSNT.....REMEMGKIRGTGCGYSGPSL 647

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10086464/runat\_11052004\_121815\_7853/app query.fasta\_1.839  
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10086464\_QCGN\_1\_1\_4237 @runat\_11052004\_121815\_7853 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mus:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1092.5	31.6	1084	12	BP184771
2	1090.5	31.6	824	14	CD435142
3	1086	31.5	788	14	CF436437
4	1071	31.0	789	14	CF436655
5	1069	31.0	759	12	BM408099
6	1066	30.9	785	14	CF436122
7	1065	30.8	757	12	BG596561
8	1063	30.8	731	14	CB655196
9	1047	30.3	692	12	BG441204
10	1030	29.8	723	14	CD839231
11	1027.5	29.8	1053	11	AY108241
12	1025	29.7	1016	11	AY108243
13	1018	29.5	666	12	BM358715
14	1013	29.3	695	13	CA072174
15	983	28.5	715	13	BU099573
16	961	27.8	770	13	BQ999193
17	931	27.0	611	14	CA237156
c 18	931	27.0	859	29	CC725849
19	929.5	26.9	809	14	CA765135
20	928	26.9	850	29	CC668754
c 21	923.5	26.7	769	14	CK283399
22	921	26.7	618	14	CA269355
23	921	26.7	651	14	CF478389
24	920	26.6	673	14	CA164704
c 25	919.5	26.6	938	29	CG333846
c 26	917.5	26.6	949	29	CG436431
27	912	26.4	679	13	BQ404121
28	904.5	26.2	674	14	CA298046
29	903	26.2	595	14	CF015663
30	903	26.2	637	13	BQ134241
31	903	26.2	645	13	BQ506869
c 32	897.5	26.0	899	14	CK096498
33	895.5	25.9	682	13	CA095337
34	895	25.9	641	14	CA248316
c 35	895	25.9	725	29	CG450877
c 36	893.5	25.9	803	14	CF243440
37	891	25.8	646	13	BQ240617
38	891	25.8	652	13	BQ849683
39	891	25.8	666	13	BU499754
40	891	25.8	691	14	CA237152
41	889	25.7	626	13	BQ582873
42	888.5	25.7	593	10	BF176907
43	887	25.7	622	14	CA999610
44	887	25.7	673	13	BQ410602
45	886.5	25.7	573	9	AV551753

ALIGNMENTS

RESULT 1  
BP184771  
LOCUS  
DEFINITION BP184771 pNS rice panicle cDNA, germ cell generating stage Oryza sativa (japonica cultivar-group) cDNA, mRNA sequence.  
ACCESSION BP184771  
VERSION BP184771.1 GI:32948199  
KEYWORDS EST.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

```

REFERENCE 1 (bases 1 to 1084)
AUTHORS Moriguchi,K., Ito,Y., Yamazaki,Y. and Kurata,N.
TITLE Finding of various plant nuclear proteins using yeast nuclear
JOURNAL transportation trap system - a proteomal approach
COMMENT Unpublished (2003)
Contact: Kazuki Moriguchi
Plant Genetics
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411-8540, Japan
Tel: 81-55-981-6872
Fax: 81-55-981-6879
Email: kmoriguc@lab.nig.ac.jp
cDNA clone obtained from nuclear transportation trap system
encoding a protein similar to Oryza sativa (japonica
cultivar-group) putative receptor protein kinase PERK1.
Location/Qualifiers
FEATURES
source
1..1084
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39947"
/tissue_type="panicle"
/dev_stage="germ cell generating stage"
/clone_lib="pNS rice panicle cDNA, germ cell generating
stage"

ORIGIN
Alignment Scores:
Pred. No.: 4,98e-26 Length: 1084
Score: 1092.50 Matches: 212
Percent Similarity: 82.88% Conservative: 30
Best Local Similarity: 72.60% Mismatches: 43
Query Match: 31.64% Indels: 7
DB: 12 Gaps: 3

US-10-086-464-2 (1-647) x BP184771 (1-1084)

Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 2 TTGCACGGAAAGGCCGACCAACATGGAGTGGCCCAACAGACTAAAGATTGCTTTGGGA 61
Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 62 GCTGCAAGGGTTTAGCTTATCTTCATGAAGACTGCCATCTCTAGATCATCTCCATCGTAT 121
Qy 400 IleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGly 419
Db 122 ATTAAGCGCTCAACATCTCTTGAATTTTAAAGTTTGAATCTAAGTTGCTGATTTTGA 181
Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 182 CTTCGTAAGTTTCCACGAGTGATAATAACACTCATGTTCGACAAGAGTAAATGGCACTTTT 241
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 242 GGATATCTAGACACAGAGTAGCACTCTCTGGCAAGCTCATCTGAGAAATCAGATGCTCTTC 301
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAlaAsnAsn 479
Db 302 TCCTATGGAGTTATGCTCTTGGTTAATAACTGCTGTCGCCAGTTGTATACAAAGTCAA 361
Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 362 ACATATATGATGATGACAGCTTGGTGTACCTGGCAAGGCCCTTTACTGATGCAAGCACTTGA 421
Qy 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGlu 519
Db 422 AATGGTAACACAGGAGCTTAGTAGCTCTCGGCTGGGAAGGATTTCAATCCCAATGAG 481
Qy 520 MetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgPro 539
Db 482 ATGGCGAAGATGATGCTTGTGCAGCTGCATGCGTACGCCATTCGGCTCGCTCGTCCGCCA 541
Qy 540 ArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGlu 559

```

---

```

Db 542 CGCATGACGACAGGTGTCCGGGCTTTGGAAGTGACGTGCTTTGGAGGATCTTAATGAA 601
Qy 560 GlyMetArgProGlyGlnSerAsnValTyrSerTyrGlyGlySerThrAspTyrAsp 579
Db 602 GGTGTTCCGGCTGTGTACAGCCGCTATTTTGGATCGGTAC---AGCAGCTCTGACTATGAT 658
Qy 580 SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGlu 599
Db 659 TCTGCCCAATACACAGGAGCATGAGAAGTTTCAGGAAGATGGCTTTT---ACCAACAAT 715
Qy 600 TyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGly 619
Db 716 AATGATACGACGACGTCAATACAGCGCACCAACACGAGGATGAGCCAGATACCTCTGCA 775
Qy 620 SerSerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly 639
Db 776 TCACGACGAGGAGGCCCAACAAACCCAGGAAGTCGAGAGCAGGACACCAAGAGAGGCGGC 835
Qy 640 -----GlnGlyTyrSerGlyProSer 646
Db 836 TACAGTGGCTACAGCTCAGGATACAGCGGAGCCTCA 871

RESULT 2
CD435142 LOCUS CD435142 824 bp mRNA linear EST 03-JUN-2003
DEFINITION EL01N0355D03.b Endosperm_3 Zea mays cDNA, mRNA sequence.
ACCESSION CD435142
VERSION CD435142.1 GI:31350785
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 824)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
Messing,J.
Sequencing of the maize endosperm ESTs
Unpublished (2002)
Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3.
FEATURES
source
Location/Qualifiers
1..824
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/tissue_type="Endosperm of 7-23DAP"
/clone_lib="Endosperm_3"
/note="vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Alignment Scores:
Pred. No.: 4.54e-26 Length: 824
Score: 1090.50 Matches: 206
Percent Similarity: 84.12% Conservative: 27
Best Local Similarity: 74.37% Mismatches: 41
Query Match: 31.58% Indels: 3
DB: 14 Gaps: 2

US-10-086-464-2 (1-647) x CD435142 (1-824)
Qy 334 LeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsn 353
Db 2 TTGTTGGCTATTGCAATTTCTGGAGGACCTTGTGCTTGTCTATGATGATTTGTCCCAAT 61

```

```

Qy 354 AsnAsnLeuGluLeuHisLeuHisGlyGluArgProThrMetGluTrpSerThrArg 373
Db 62 AACACATGGATTCACATACGCGAAGGTGACCACTATGGAGTGGCCCTGCTAGA 121
Qy 374 LeuValLeuAlaLeuGlySerAlaGlyLeuSerTyLeuHisGluAspCysAsnPro 393
Db 122 TTAAGATCAGTTGGGTGCTGCCAAGGTTAGCTTATCTTCATGAAGCTGCATCCA 181
Qy 394 LysIleIleHisArgAspIleLeuAlaSerAsnIleLeuIleAspPheLeuPheGluAla 413
Db 182 AAGATCATCCATCGTGCATGAAGGATCTAACATCTCTTGACTTCCATTTGAAGCT 241
Qy 414 LysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThr 433
Db 242 AAGGTGCTGATTTGGACTTGCAGTTCTACTACTGATTAACAACACCCATGTTTCACA 301
Qy 434 ArgValMetGlyThrPheGlyTyLeuAlaProGluTyAlaAlaSerGlyLeuLeuThr 453
Db 302 AGAGTAATGGGCACCTTTGGGTATTTGGCACCTGAGTATGATCTCTTGGCAAGCTAACA 361
Qy 454 GluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyValArg 473
Db 362 GAAAAATCCGATGATTTCTTCGAGTCAATGCTCTTGAGCTTATTTACTTGGCGCGCA 421
Qy 474 ProValAspAlaAsnValTyValAspAspSerLeuValAspTTPAlaAatGProLeu 493
Db 422 CCAGTTGACACCAACCAACATATATGGATGACAGCTTGGTTGACTGGGCAGGCCATTA 481
Qy 494 LeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsn 513
Db 482 CTGATGCGAGCACCTTGAGGATGGTGAATGATTTTATGATCTCTTGGGCTGCGGAAAG 541
Qy 514 GlyTyAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHis 533
Db 542 GACTTCAATCCTAATAGATGGCAAGATGATACCTGTGTCAGCTGATGTGTACGCCAT 601
Qy 534 SerAlaArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSer 553
Db 602 TCTGCAGCTGCTGGCCACGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
Qy 554 LeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTySerSerTyGly 573
Db 662 TTGGAGGACCTTAAATGAAGGTGTTTGGCCTGGCCATAGCCGCTTCTTTGGGTCTATAC 718
Qy 574 GlySerThrAspTyAspSerSerGlnTyAsnGluAspMetLysLysPheArgLysMet 593
Db 719 AGCAGCTCCGATTAACGATCTTGGCCAGCTACCAACGAGGACATGAAGAAGTTCAAGAAGATG 778
Qy 594 AlaLeuGlyThrGlnGluTyAsnAlaThrGlyGluTySerAsnProThr 610
Db 779 GCATTCAACAACAAC-----TATACCAGCAGCCCATACAGCGGCCCAACC 823

```

## RESULT 3

CF436437

LOCUS

DEFINITION EST672782 normalized cDNA library of onion Allium cepa cDNA clone

ACACM44, mRNA sequence.

ACCESSION CF436437

VERSION CF436437.1

KEYWORDS GI:34459127

SOURCE EST.

ORGANISM Allium cepa (onion)

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae; Allium.

## REFERENCE

AUTHORS

TITLE

Expressed Sequence Tags from a normalized library of mixed onion

JOURNAL

COMMENT

Unpublished (2003)  
Contact: Hovey MJ  
Department of Horticulture

USDA-ARS and University of Wisconsin  
1575 Linden Drive, Madison, WI 53706, USA  
Tel: 608-262-1830  
Fax: 608-262-4743  
Email: mjhavey@facstaff.wisc.edu  
TIGR sequence name ACACM44TR. For more information:  
http://haveylab.hort.wisc.edu  
Seq primer: CAG GAA ACA GCT ATG ACC.

## FEATURES

source

Location/Qualifiers

1..788

/organism="Allium cepa"

/mol\_type="mRNA"

/cultivar="Red Creole (bulbs), unknown (callus), Ebano &amp;

Texas Legend (roots)"

/db\_xref="taxon:4679"

/clone="ACACM44"

/tissue\_type="Callus, roots, and young bulbs"

/clone\_lib="normalized cDNA library of onion"

/notes="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:

EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA

from callus, roots, and young bulbs were combined to

synthesize the library. Normalized by proprietary

low-copy transcripts was performed by proprietary

techniques of Invitrogen."

## ORIGIN

Alignment Scores:  
Pred. No.: 6,07e-26 Length: 788  
Score: 1086.00 Matches: 204  
Percent Similarity: 87.79% Conservative: 26  
Best Local Similarity: 77.86% Mismatches: 32  
Query Match: 31.45% Indels: 0  
DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x CF436437 (1-788)

```

Qy 260 LysSerThrPheThrTyGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
Db 2 AAGAGCATTTTCATATGATGAGATTTGGCCATAGCTACAAATGGTTTTCGACTATAT 61
Qy 280 LeuLeuGlyGlnGlyGlyPheGlyTyValHisGlyValLeuProSerGlyLysGlu 299
Db 62 CTCTTGGGCAAGTGGATTTGGATATGTGCACAAAGAGTACTTCCAAACGGTAAAGAA 121
Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
Db 122 GTGGCTATCAAGCAGTTGAAAGCCGGAAGCGGAGCGAGGTGAGTTTCAAGCAGAG 181
Qy 320 ValGluIleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyCysIle 339
Db 182 GTTGAATATATCATGTCGATGCATCATAGGCATTTGGTTCTTTAGTTGGCTATTGCATT 241
Qy 340 AlaGlyAlaLysArgLeuLeuValTyGluPheValProAsnAsnLeuGluLeuHis 359
Db 242 TCAGGAGATCATAGATTGCTTGTATGAATATGTTCTTAATAAAACCTTGAGTTCCAT 301
Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 302 TTGCATGAAAAGATGTTTCCACTGATGTTGGCCAAACGGGTGAAATTTGCTTTGGGT 361
Qy 380 SerAlaLysGlyLeuSerTyLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 362 TCTGCCAAGGGTGGCATATCTCCATGAAGATTGTCATCCCAAAATTTATTCATCGTAT 421
Qy 400 IleLysAlaSerAsnIleLeuIleAspPheGlyPheGluAlaLysValAlaAspPheGly 419
Db 422 ATTAAGCAGCAAAATATCTTCTTGATGATCAACTTTGAGGCTAAGGTTGCAGATTTTGGC 481
Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 482 CTTGCAAAAGTTTCACTGTGTTGAAAATAACACCCATGTTTCTACAGAGTCATGGGAACATTT 541
Qy 440 GlyTyLeuAlaProGluTyAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459

```

```

|||||GGGATATCGGCACAGATATGCAATCTCTGGTAACTAAAGTAAATCAGATGTCCTTT 601
Db 542 GGGATATCGGCACAGATATGCAATCTCTGGTAACTAAAGTAAATCAGATGTCCTTT 601

|||||SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 602 TCATTGGAGTCATGCTTTTGGAGCTGATTACTGGAAAGCCAGCCAGTTGATGCAACCAA 661

|||||ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 662 ACTTATACAGATGATAGTTTGGTCGATTGGGCAAGACCATTGCTGCAGAAAGGCTATGGAG 721

|||||GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Qy 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db 722 CATGATGATTATGATGAGCTCATTGACAGAAGGCTTAGGAGATAAATCAACCATGACGAA 781

520 MetAla 521
782 ATGGCA 787

CF436655 789 bp mRNA linear EST 04-SEP-2003
LOCUS EST673000 normalized cDNA library of onion Allium cepa cDNA clone
DEFINITION ACACK26, mRNA sequence.
ACCESSION CF436655
VERSION EST.
KEYWORDS Allium cepa (onion)
SOURCE Allium cepa
ORGANISM Allium cepa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
Allium.
REFERENCE 1 (bases 1 to 789)
AUTHORS Havey,M.J., Cheung,F., Van Aken,S., Utterback,T. and Town,C.D.
TITLE Expressed Sequence Tags from a normalized library of mixed onion
tissues (Allium cepa)
JOURNAL Unpublished (2003)
COMMENT Contact: Havey MJ
Department of Horticulture
USDA-ARS and University of Wisconsin
1575 Linden Drive, Madison, WI 53706, USA
Tel: 608-262-1830
Fax: 608-262-4743
Email: mjhavey@facstaff.wisc.edu
TIGR sequence name ACACK26TR. For more information:
http://haveylab.hort.wisc.edu
Seq primer: CAG GAA ACA GCT ATG ACC.
FEATURES
source
Location/Qualifiers
1..789
/organism="Allium cepa"
/mol_type="mRNA"
/cultivar="Red Creole (bulbs), unknown (callus), Ebano &
Texas Legend (roots)"
/db_xref="taxon:4679"
/clone="ACACK26"
/tissue_type="Callus, roots, and young bulbs"
/clone_lib="normalized cDNA library of onion"
/note="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1:
EcoRI (5'); Site 2: NotI (3'); Equal molar amounts of mRNA
from callus, roots, and young bulbs were combined to
synthesize the library. Normalization to enrich for
low-copy transcripts was performed by proprietary
techniques of Invitrogen."

ORIGIN
Alignment Scores:
Pred. No.: 1,82e-25 Length: 789
Score: 1071.00 Matches: 203
Percent Similarity: 86.64% Conservative: 24
Best Local Similarity: 77.48% Mismatches: 35
Query Match: 31.02% Indels: 0
DB: 14 Gaps: 0

```

```

US-10-086-464-2 (1-647) x CF436655 (1-789)

Qy 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
Db 2 AAGAGCACTTTCACATATGAAGAATTGGCGATAGCTACAAATGGGTTTTCGCACTATAAT 61

Qy 280 LeuLeuGlnGlnGlyPheGlyTyrValHisGlyValLeuProSerGlyValSglu 299
Db 62 CTTCTTGGCGAAGGTGGATTGGATATGTGCACAAAGGAGTACTTCCAAACGGTAAGAA 121

Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
Db 122 GTGGCTATCAAGCAGTTGAAAGCCGGAAGCGGACAGGGGAGCGTGAGTTTCAGCAGAG 181

Qy 320 ValGluIleLeuSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIle 339
Db 182 GTTGAATTTATCAGTCGAGTCATCATAGGCATTTGGTTCTTTAGTTGGCTATTGCATT 241

Qy 340 AlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuLeuHis 359
Db 242 TCAGGAGATCATAGATTGCTTGTCTATGAATATTTCTTAATAAACCCCTTGAGTTCAT 301

Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 302 TTGCATGGAAAGATGTTCCACCTATGATGGCCAAACGGGTTAAAAATTGCTTGGGT 361

Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 362 TCTGCCAAGGGGTTGGCATATCTCATGAAGATTGTCTCCAAAATTTATTCATCGTGAT 421

Qy 400 IleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGly 419
Db 422 ATTAAGACAGCAATAATTTCTTTGAGATGAACATTTGAGGCTAAGGTTGAGATTTTGGC 481

Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 482 CTTGCAAAAGTTCACCTGTTGAAATAACACCCATGTTCTACACGAGTCATCGGAACATTT 541

Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 542 GGGTATCTGGCACCAGAAATATGCATCTCTGTGTAACCTAAGTGATAAATCAGATGTCCTTT 601

Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 602 TCATTGGAGTCATGCTTTTGGAGCTGATTACTGGAAGACGACCATGTTGATGCAACCCAA 661

Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGlu 499
Db 662 ACTTATACAGATGATAGTTTGGTCGATTGGCCAGAACCATTTGCTGCAGAAAGGCTATGGAG 721

Qy 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGlu 519
Db 722 CATGATGATTATGATGAGCTCATTGACAGAAGGCTTAGGAGATAAATACCCCATGACGA 781

520 MetAla 521
782 ATGGCA 787

RESULT 5
LOCUS BM408099 759 bp mRNA linear EST 10-MAR-2003
DEFINITION ES1582426 potato roots Solanum tuberosum cDNA clone CPRO33021 5'
end, mRNA sequence.
ACCESSION BM408099
VERSION BM408099.1 GI:18259729
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 759)
AUTHORS van der Hoeven,R., Sun,H., Karamycheva,S.A., Tsai,J., Van Aken,S.,

```

Utterback, T., Chiening, A., Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.  
Generation of ESTs from potato roots  
Unpublished (2001)  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:  
http://genome.arizona.edu/orders/  
Seq primer: T3.

## FEATURES

Location/Qualifiers  
1..759  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="cPRO33021"  
/tissue\_type="roots"  
/dev\_stage="in vitro grown stem cuttings"  
/lab\_host="SOLR"  
/clone\_lib="potato roots"  
/notes="vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley Lab; sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

## ORIGIN

Alignment Scores:  
Pred. No.: 2,04e-25 Length: 759  
Score: 1069.00 Matches: 200  
Percent Similarity: 89.29% Conservative: 25  
Best Local Similarity: 79.37% Mismatches: 27  
Query Match: 30.96% Indels: 0  
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BM408099 (1-759)

Qy 312 GlyGluArgGluPheGlnAlaGluValGluIleSerArgValHisArgHisLeu 331  
Db 2 GGGGAACTGTAATTCAGGCGGAAGTTGAGATTATTAGCGAGTACATCAAGCATCTT 61  
Qy 332 ValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuValTyrGluPheVal 351  
Db 62 GTGTCTCTGTGGATACCTGCATTACTGGGCTCAGAGACTGCTTGTATTGAGTTTGT 121  
Qy 352 ProAsnAsnLeuLeuGluHisLeuHisGlyGluArgProThrMetGluTrpSer 371  
Db 122 CCAACAATATCTTGGAAATTCATTACAGGAAGGAAGGAGCTCTCTGGATTGCCA 181  
Qy 372 ThrArgLeuValIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCys 391  
Db 182 ATACGGCTAAAGATTGCTTAGGCTCAGCTAAAGGACTGGCATATCTGCATGAAGACTGC 241  
Qy 392 AsnProLysIleLeHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPhe 411  
Db 242 CAACCGAAATCATCTACCGGTGATCATCAAGGAGCTAATATCTATCGACTTTAAATTT 301  
Qy 412 GluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrHisVal 431  
Db 302 GAGCTAAGGTGCTGATTGGACTTGGCCANGCTAACTCTCTGATGTAATATCATCATGTC 361  
Qy 432 SerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLys 451  
Db 362 TCCACCAAGATGATGGGAACCTTTGGGTAATTGGCTCCAGAAATATGCTTCTCTGGAAG 421  
Qy 452 LeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGly 471  
Db 422 CTTACAGACAAGTCAGACGATTTCTCTTGGTGAATGCTTCTTGATGTTGATAACTGGA 481

Qy 472 ArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTrpAlaArg 491  
Db 482 CGTCGGCTCTGTGACTCTACTCAATCATCATCGAAGATAGTTTGGTGGAGCTGGCAGCT 541  
Qy 492 ProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMet 511  
Db 542 CCATTACTCACAGAGCTTTAGAAAGATGAAAGTTTGTATACCTCTGTTGATCGTCGGCTA 601  
Qy 512 AsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAlaCysVal 531  
Db 602 GAAATGATTATACCATATAGATGGCTCGCATGGTTGCTTGTCTGCTGCTGCTG 661  
Qy 532 ArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsn 551  
Db 662 CGTCATTTCAGCAAGCGTAGACCAAGCAATGACACAGGTTCTCCGAGCTTGGAAAGGAGAT 721  
Qy 552 ValSerLeuSerAspLeuAsnGluGlyMetArgPro 563  
Db 722 GTCTCATTTACAGACCTTTACGAAGGGATTAACCT 757

## RESULT 6

CF436122 785 bp mRNA linear EST 04-SEP-2003  
LOCUS EST672467 normalized cDNA library of onion Allium cepa cDNA clone  
DEFINITION ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

ACACJ26, mRNA sequence.

## ORIGIN

Alignment Scores:  
Pred. No.: 2,61e-25 Length: 785  
Score: 1066.00 Matches: 204  
Percent Similarity: 86.64% Conservative: 23  
Best Local Similarity: 77.86% Mismatches: 34  
Query Match: 30.87% Indels: 1  
DB: 14 Gaps: 0

1..785  
/organism="Allium cepa"  
/mol\_type="mRNA"  
/cultivar="Red Creole (bulbs), unknown (callus), Ebano & Texas Legend (roots)"  
/db\_xref="taxon:4679"  
/clone="ACACJ26"  
/tissue\_type="Callus, roots, and young bulbs"  
/clone\_lib="normalized cDNA library of onion"  
/notes="Vector: pCMVSPORT6.1-ccdb (Invitrogen); Site 1: EcoRV (5'); Site 2: NotI (3'); Equal molar amounts of mRNA from callus, roots, and young bulbs were combined to synthesize the library. Normalization to enrich for low-copy transcripts was performed by proprietary techniques of Invitrogen."

Email: mjhavey@facstaff.wisc.edu

TIGR sequence name ACACJ26TR. For more information:

http://haveyab.hort.wisc.edu

Seq primer: CAG GAA ACA GCT ATG ACC.

Location/Qualifiers

source

```
US-10-086-464-2 (1-647) x CF436122 (1-785)

Qy 260 LysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsn 279
Db 2 AAGAGCACTTTCACATATGAAGATTGGCGATAGCTACAAATGGTTTCCGACTATAAT 61

Qy 280 LeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGlu 299
Db 62 CTTCTTGGCGAAGTGGATTGGATATGTGCACAAAGGAGTACTTCCAAACGGGTAAAGAA 121

Qy 300 ValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGlu 319
Db 122 GTGGCTATCAGCAGAGTTGAAGCCGAGCGGACAGCGGAGCGTGAGTTTCAAGCAGAG 181

Qy 320 ValGluLeuLeuSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIle 339
Db 182 GTTCAAAATTATCAGTCGAGTCATCATAGGCATTTGGTTCTTTAGTTGGCTATTGCATT 241

Qy 340 AlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHis 359
Db 242 TCAGGAGATCATAGATTGCTTGTCTATGAATATGTTCTTAATAAACCCTTGAGTTCCAT 301

Qy 360 LeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGly 379
Db 302 TTGCATGGAAAGATGTTCCACCTATGGAATGGCCACCGGTTAAATAATGCTTTGGGT 361

Qy 380 SerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAsp 399
Db 362 TCTGCCAAGGGTTGGCATATCTCCATGAGATTGTCTATCCAAATAATTATTCATCGTAT 421

Qy 400 IleLysAlaSerAsnIleLeuLysPheLysPheGluAlaLysValAlaAspPheGly 419
Db 422 ATTAAGCAGCAATAATTCTTCTGAGATGAACATTTGAGGCTAAGGTTGCAGATTTTGGC 481

Qy 420 LeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 482 CTTCCAAAGTTCACCTGTTGGAATAAACACCATGTTTCTACACGAGTCATGGGAACATT 541

Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 542 GGGTATCTGGCCACGAGATATGCATCTCTGTGTAACCTAAGTGATAAATCAGATGTCTTT 601

Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsn 479
Db 602 TCATTGAGTGCATGCTTTTGGAGCTGATTACTGGAAAGCGACCATGTTGATGCAACCCAA 661

Qy 480 ValTyrValAspAspSer-LeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerG 499
Db 662 ACTTATACAGATGATAGTTTGGTGGATTTGGGCAAGACCATGCTGCAGAAGGCTATGGA 721

Qy 499 uGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGlu 519
Db 722 GCATGATGATTATCATGAGCTCATTTGACAGAAAGGCTAGGAGATAATACACCATGACA 781

Qy 519 uMet 520
Db 782 AATG 785

RESULT 7
LOCUS BG596561
DEFINITION EST495239 cSTS Solanum tuberosum cDNA clone cSTS15A23 5' sequence, mRNA sequence.
ACCESSION BG596561
VERSION BG596561.1 GI:13614701
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 757)

AUTHORS van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A.,
Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13P-R.
FEATURES
source Location/Qualifiers
1..757
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS15A23"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/clone_lib="cSTS"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
ORIGIN
Alignment Scores: 2.72e-25 Length: 757
Pred. No.: 1065.00 Matches: 200
Score: 89.16% Conservative: 22
Percent Similarity: 30.32% Mismatches: 27
Best Local Similarity: 80.84% Indels: 0
Query Match: 12 Gaps: 0
DB: 0
US-10-086-464-2 (1-647) x BG596561 (1-757)
Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db 1 GCAGTTAAACAGCTTAAGGCTGGAGTGGACACAGGGGACGTGAATTCAGCGCGAAGTT 60
Qy 321 GluIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db 61 GAGATTATTAGCCGAGTACATCAACGATCTTGTGTCTCTTGTGGATACTGCATTACT 120
Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeu 360
Db 121 GGGGCTCAGAGCTGCTTGTATTAGTGTGTTTCCAAACCAATACTTTTGGAAATTTCAATTA 180
Qy 361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db 181 CACGAAAGGGAAGGCGCTCTTGGATTGGCCATACGGCTAAAGATGTCTTAGGCTCA 240
Qy 381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db 241 GCTAAAGGACTGGCATATCTGCATGAAGACTGCCAACCGAAATAATCAATCCCGTGATATC 300
Qy 401 LysAlaSerAsnIleLeuLysPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db 301 AAGGACGCTAATACTATCTATCGACTTTAAATTTTTCAGGCTAAGGTTGCTGATTTTGGACTT 360
Qy 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db 361 GCCAAGCTAACTTCTGATGTTAATACTCATGTCTCCACAGAGTGATGGGAACCTTTGG 420
Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db 421 TATTGGCTCCAGAAATATGCTTCTCTGGAAAGCTTACAGACAAGTGCAGAGCTATTCTCC 480
```



of the cotton fiber  
Unpublished (2000)  
Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAATACGACTCACTATAGG  
High quality sequence stop: 690.

FEATURES  
source  
1..692  
/organism="Gossypium arboreum"  
/mol\_type="mRNA"  
/strain="AKA"  
/cultivar="8400"  
/db\_xref="taxon:29729"  
/clone="GA\_Ea0012C15f"  
/tissue\_type="Fibers isolated from bolls harvested 7-10 dpa"  
/lab\_host="E. coli"  
/clone\_lib="Gossypium arboreum 7-10 dpa fiber library"  
/note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
Alignment Scores:  
Pred. No.: 9.38e-25 Length: 692  
Score: 1047.00 Matches: 198  
Percent Similarity: 94.71% Conservat: 17  
Best Local Similarity: 87.22% Mismatches: 12  
Query Match: 30.32% Indels: 0  
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BG441204 (1-692)

Qy 247 LeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrGlu 266  
Db 11 CTTCCGCTGGTACCTCGATGATTTCTTAGTTCTCGAAGACATTTTAGCTATGAA 70

Qy 267 GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPhe 286  
Db 71 GAATTAGCGAGCAACGAGTGGCTTCCTCGGAAGTTAACTTCTTGGACAAGTGGTTT 130

Qy 287 GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306  
Db 131 GGGTACGTACACAAAGGAGTTCTCCCTAATGGGAAGGAGTAGCAGTAAAGCAACTCAAG 190

Qy 307 ValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgVal 326  
Db 191 GCTGGAATGGCGAAGCGGAGAGAAATTCAGGCTGAAGTTGAGATCATTTAGCCGCTC 250

Qy 327 HisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeu 346  
Db 251 CATCACAAACATCTGCTCTCATTTGTCGATCTGATTTCTGGACAATTAAGATGCTT 310

Qy 347 ValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGluGlyArgPro 366  
Db 311 GTTTATGAGTTGTTTCCAAACAACACCTTGGAGTTTCACTTGCATGGGAAGGCGGCTG 370

Qy 367 ThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyr 386  
Db 371 ACCATGGAATGGCCGACGAAGGATGAAATTCCTTTAGGATCTGCAAAAGGACTGGCATAT 430

Qy 387 LeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 406  
Db 431 CTTTATGAAGATTGTCATCTCAAGATCATTCACCGTGATTAAGGCGCTAATATCTG 490

Qy 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426  
Db 491 TTGGATTTCAAGTTTGAAGCAAGGTTGCTGATTTTGGACTAGCGAAATTTGCTTCGAT 550

Qy 427 ThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyr 446

Db 551 GTCAACACGACGCTCTCCACACGCGTATGGGTACTTTCCGGTATTTTAGCCCTGAGTAT 610

Qy 447 AlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeu 466  
Db 611 GCTTCAAGTGAAGCTCACTGATAATCAGATGTTTCTCTCTCGGGGTATGCTTTTG 670

Qy 467 GluLeuIleThrGlyArgArg 473

Db 671 GAGTTGATTACCGGTACAG 691

RESULT 10  
CD839231 723 bp mRNA linear EST 10-JUL-2003  
LOCUS RFO2.114104F010529 RFO2 Brassica napus cDNA clone RFO2114104, mRNA  
DEFINITION sequence.  
ACCESSION CD839231  
VERSION CD839231.1 GI:32521171  
KEYWORDS EST.  
SOURCE Brassica napus (rape)  
ORGANISM Brassica napus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 723)  
AUTHORS Genoplatte.  
TITLE Genoplatte, a major partnership french program in plant genomics  
JOURNAL Unpublished (2003)  
COMMENT Contact: Genoplatte  
Genoplatte  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplatte' (<http://www.genoplatte.com>  
and <http://genoplatte-info.infobiogen.fr>).

FEATURES  
source  
1..723  
/organism="Brassica napus"  
/mol\_type="mRNA"  
/cultivar="samourai (restored line)"  
/db\_xref="taxon:3708"  
/clone="RFO2114104"  
/tissue\_type="anthers"  
/clone\_lib="RFO2"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.37e-24 Length: 723  
Score: 1030.00 Matches: 199  
Percent Similarity: 100.00% Conservat: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 29.83% Indels: 0  
DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x CD839231 (1-723)

Qy 449 SerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeu 468  
Db 2 AGCGGAAAGCTCACCGAGAAGTCTGACGTTTCTCATTTGGCGTTTGTGAGGCTC 61

Qy 469 IleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAsp 488  
Db 62 ATTACCGGCGCTCGACCGTTGATGCCAACAAATGCTATGTAGATGACAGCTTAGTGAC 121

Qy 489 TrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGlyLeuAlaAsp 508  
Db 122 TGGGACGACCATTTGCTTAACCGACATCTCAGCAAGGAGACTTTGAGGTTTAGCTCAT 181

Qy 509 AlaLysMetAsnAsnGlyTyrAspArgGluMetAlaArgMetValAlaCysAlaAla 528  
Db 182 GCAAGATGAATAATGGGTATGACAGAGAGAGATGGTCGCATGGTTGTTGCTGCTCG 241

Qy 529 AlaCysValArgHisSerAlaArgProArgMetSerGlnIleValArgAlaLeu 548  
 |||||  
 Db 242 GCTTGTGTCGCCATTGAGCTCGCGAGACCTCGCATGAGCCAGATTGTGCGTGCCTTA 301  
 |||||  
 Qy 549 GluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnVal 568  
 |||||  
 Db 302 GAAGGAAATGATCATTGTCAGATCTTAACAAAGGAGTACAGACAGGTCAAGCAATGTGA 361  
 |||||  
 Qy 569 TyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLys 588  
 |||||  
 Db 362 TACAGCTCATACGAGGAGGACCCGATATGACTCGAGCCAGTACATGAGACATGAG 421  
 |||||  
 Qy 589 LysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsn 608  
 |||||  
 Db 422 AAGTTTAGGAAATGGCTCTTTGGAACCTCAAGAGTACAAAGCCAGCGGTGAGTACAGTAAT 481  
 |||||  
 Qy 609 ProThrSerAspTyrGlyLeuTyrProSerGlySerSerGlyGlnThrThrArg 628  
 |||||  
 Db 482 CCGACCAAGCGACTATGACTGTACCCGCTCGTTCAGAGCAGCGGCGCAACACACAGC 541  
 |||||  
 Qy 629 GluMetGluMetGlyLysIleLysArgThrGlyGlnGlyTyrSerGlyProSerLeu 647  
 |||||  
 Db 542 GAAATGGAGATGGGGAAGATTAGAGAAACCGGTGAGGGTTATAGTGAACCTTCTT 598  
 |||||

## RESULT 11

AY108241  
 LOCUS AY108241 1053 bp mRNA linear HTC 16-OCT-2002  
 DEFINITION Zea mays PC0134818 mRNA sequence.  
 ACCESSION AY108241  
 VERSION AY108241.1 GI:21211319  
 KEYWORDS HTC.

## SOURCE

ORGANISM Zea mays

Zeas mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 1053)  
 Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

## REFERENCE

AUTHORS

Maize Mapping Project/DuPont Consensus Sequences for Design of

## TITLE

Overgo Probes

Unpublished (2002)

2 (bases 1 to 1053)

Coe,E.H.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

## FEATURES

source

1..1053  
 Location/Qualifiers  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="MaizeDB:637889"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

## ORIGIN

Alignment Scores: 5.61e-24 Length: 1053  
 Pred. No.: 1027.50 Matches: 203  
 Score:

Percent Similarity: 79.93% Conservative: 28  
 Best Local Similarity: 70.24% Mismatches: 46  
 Query Match: 29.76% Indels: 12  
 Db: 11 Gaps: 4  
 US-10-086-464-2 (1-647) x AY108241 (1-1053)  
 Qy 365 ArgProThrMetGluTrpSerThrArgLeuLysIle-AlaLeuGlySerAlaLysGlyLe 384  
 |||||  
 Db 6 CGTCCG-----GAGTGGCCTGTAGATTAAGATCATGTTTGGGTCTGCTCAAGGGTTT 59  
 |||||  
 Qy 384 uSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAs 404  
 |||||  
 Db 60 AGCTTATCTTCATGAAGACTGCCATCCAAAGATCATCCATCGTGACATAAAGGCATCTAA 119  
 |||||  
 Qy 404 nIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAl 424  
 |||||  
 Db 120 CATCTCTTCGTGACTTCCTCAATTTGAAGCTAAGTTGCTGATTTGGATTCGCAAGTTCCAC 179  
 |||||  
 Qy 424 aSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAlaPr 444  
 |||||  
 Db 180 TACTGATAACAACACCCCATGTTTCGACACAGAGTAATGGGCACCTTTGGGTATTGGCACC 239  
 |||||  
 Qy 444 oGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValVa 464  
 |||||  
 Db 240 TGAGTATGCATCTCTCTGGCAAGCTAA CAGAAAAATCCGATGTATTTCCTTCGGAGTCAT 299  
 |||||  
 Qy 464 lLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnValTyrValAspAs 484  
 |||||  
 Db 300 GCTTCTTGAGCTTATTACTGGGCGGCGCACCAGTTGACACCAACCATATATATGGATGA 359  
 |||||  
 Qy 484 pSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGl 504  
 |||||  
 Db 360 CAGCTTGGTTGACTGGGCAAGGCCATTACTGATCGGAGCACCTTGAGGATGGTGAATATGA 419  
 |||||  
 Qy 504 uGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMetAlaArgMetVa 524  
 |||||  
 Db 420 TGCTTTAGTGCATCTCTCGGCTGGGAAAGGACTTCAATCCTAATGAGATGGCAAGATGAT 479  
 |||||  
 Qy 524 lAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgProArgMetSerGlnI 544  
 |||||  
 Db 480 AGCCTGTGCAGCTGCATGTGTACGCCATTCTGCACGTCGTCGCCACGATAGTACAGT 539  
 |||||  
 Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564  
 |||||  
 Db 540 CGTTCGGGCTTTGGAGGGCAATGTGCTTTGGAGGACCTTAATGAAGGTGTTCGSCCTGG 599  
 |||||  
 Qy 564 yGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGlnTyrAs 584  
 |||||  
 Db 600 CCATAGCCGGCTTCTTTGGGTCTATAC---AGCAGCTCCGATTACGATTCCTGCCAGTACAA 656  
 |||||  
 Qy 584 nGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAlaThrGl 604  
 |||||  
 Db 657 CGAGGACATGAAGAAGTTCAAGAAGATGGCATTCACCAACAAAC-----TATACACGAG 710  
 |||||  
 Qy 604 yGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSerGluGl 624  
 |||||  
 Db 711 CCAATACAGCGCGCCCAACCACTGATATGACAGATACCTCTGTCATCAACAGCAGGAGGG 770  
 |||||  
 Qy 624 yGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly----- 639  
 |||||  
 Db 771 CCACCAGACGCAAGAGATGGAGTCGGGTGCAATGAAGAAAGGTGCTTACAGTGGTGGCTA 830  
 |||||  
 Qy 640 ----GlnGlyTyrSerGlyProSer 646  
 |||||  
 Db 831 CAGCTCAGGATACAGCGGAGCCCTCG 855  
 |||||  
 RESULT 12  
 AY108243  
 LOCUS AY108243 1016 bp mRNA linear HTC 16-OCT-2002  
 DEFINITION Zea mays PC0134814 mRNA sequence.  
 ACCESSION AY108243  
 VERSION AY108243.1 GI:21211321

```

KEYWORDS      HTC.
SOURCE         Zea mays
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE     1 (bases 1 to 1016)
AUTHORS       Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
               Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
               Zea Mapping Project/DuPont Consensus Sequences for Design of
               Overgo Probes
TITLE         Unpublished (2002)
JOURNAL       2 (bases 1 to 1016)
AUTHORS       Coe,E.H.
REFERENCE     Direct Submission
TITLE         Submitted (25-APR-2002) Maize Mapping Project, University of
               Missouri, Columbia, MO 65211, USA
COMMENT       If you are interested in getting corresponding physical clones,
               these are publicly available from ZmDB and may be found by BLAST
               searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
               www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
               maize cDNA sequences is either Virginia Walbot, Stanford or Pat
               Schnable, Iowa State, then clones may be requested from ZmDB:
               www.zmdb.iastate.edu.
FEATURES      Location/Qualifiers
               1..1016
                /organism="Zea mays"
                /mol_type="mRNA"
                /db_xref="taxon:4577"
                /clone_lib="Maize Mapping Project/DuPont Consensus
                Library"
                /note="this sequence is part of a project of EST
                assemblies resulting from the application of public
                contigs to seed DuPont contigs; this resource was
                assembled by DuPont as part of a collaboration for the
                overgo addressing of BACs in conjunction with the Maize
                Mapping Project"
ORIGIN
Alignment Scores:
Pred. No.:      6,53e-24      Length:      1016
Score:          1025.00      Matches:    200
Percent Similarity: 82.14%      Conservative: 30
Best Local Similarity: 71.43%      Mismatches: 42
Query Match:    29.68%      Indels:     8
DB:             11          Gaps:       4

US-10-086-464-2 (1-647) x AY108243 (1-1016)
Qy 373 ArgLeuYsIlleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsn 392
Db 3 AGATTAAAGATCGCTCTGGGTGCTGCCAAGGTTTAGCTTATCTTCATGAAGACTGCCAT 62
Qy 393 ProlysllelleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGlu 412
Db 63 CCAAAGATCATCCACCGGCATTAAGGCATCTAACATCTTCTTGACTTCATTAATTGAA 122
Qy 413 AlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSer 432
Db 123 GCTATGTTGCTGACTTTGGACTTGGCAAGTTCACCTACTACTGATATACACACCCTGTGTCA 182
Qy 433 ThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeu 452
Db 183 ACAAGAGTAATGGGCACCTTTGGGTATTTCGACCCGAGTAGTCAGCATCTGGCAAGCTC 242
Qy 453 ThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArg 472
Db 243 ACAGAAAATCCGATGTTATTTCTTCGGAGTCATGCTTCTTGAGCTTATTACTGGCGG 302
Qy 473 ArgProValAspAlaAsnValTyrValAspAspSerLeuValAspTyrAlaArgPro 492
Db 303 CGACCAATTGACACACCAACCAACATATATGATGATGACAGCTTGGTTGACTGGCGAAGGCCA 362

```

---

```

Qy 493 LeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsn 512
Db 363 TTACTGATGAGAGCGCTCGAGGATGGTCAATATGCTTTGGTGGATCTCTCGGCTCGGA 422
Qy 513 AsnGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArg 532
Db 423 AAGGACTTCAATCTCAAGAGATGGCAAGATAGCATGCTGTGGGCTCTCGGAGCGACGTG 482
Qy 533 HisSerAlaArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal 552
Db 483 CATTCGCGACGTCGTGCGCCACCAATGATGATGATGATGATGATGATGATGATGATGAT 542
Qy 553 SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr 572
Db 543 TCTTTGGAGGACCTTAATGAAGGTGTTGACCTCGCCATAGCCCTCTTTTGGGTTCATAT 602
Qy 573 GlyGlySerThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysLysPheArgLys 592
Db 603 ---AGCAGCTCTGATTACGATTCGGGCGCATACCAAGAGGACATGCGAAGTTGAGGAG 659
Qy 593 MetAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsn 612
Db 660 ACGCGTTTCAATAACAACACCTACACCCAGCAGC---CAATACAGCGCGCCCAACAGTGAA 716
Qy 613 TyrGlyLeuTyrProSerGlySerSerGluGlyGlnThrThrArgGluMet---Glu 631
Db 717 TACGCCCAGGTGCGGCTCTGGATCAAGCAGCGAGCGCGCCAGCAGCAGGAGGAGGAG 776
Qy 632 MetGlyLysIleLysArgThrGly-----GlnGlyTyrSerGlyProSer 646
Db 777 TCGGGTCGAGTGAAGAAAGGTGGCTACAGCGGCTACAGCTCCGGATACAGCGAGCCTCG 836

RESULT 13
BM358715
LOCUS      666 bp      mRNA      linear      EST 09-JAN-2002
DEFINITION GA_Ba0012D16r Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA_Ba0012D16r, mRNA sequence.
ACCESSION  BM358715
VERSION     BM358715.1
KEYWORDS    GI:18099461
SOURCE      Gossypium arboreum
            Gossypium arboreum
ORGANISM    Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE   1 (bases 1 to 666)
            Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
            Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
            An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
            Unpublished (2000)
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total High Quality bases = 646
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence stop: 666.
FEATURES    Location/Qualifiers
               1..666
                /organism="Gossypium arboreum"
                /mol_type="mRNA"
                /strain="AKA"
                /cultivar="8400"
                /db_xref="taxon:29729"
                /clone="GA_Ba0012D16r"
                /tissue_type="Fibers isolated from bolls harvested 7-10
                dpa"
                /lab_host="E. coli"

```

/clone\_lib="Gossypium arboreum 7-10 dpa fiber library"  
/note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

## ORIGIN

Alignment Scores:  
Pred. No.: 7,55e-24 Length: 666  
Score: 1018.00 Matches: 191  
Percent Similarity: 95.41% Conservative: 17  
Best Local Similarity: 87.61% Mismatches: 10  
Query Match: 29.48% Indels: 0  
DB: 12 Gaps: 0

US-10-086-464-2 (1-647) x BM358715 (1-666)

```
Qy 247 LeuProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrglu 266
|||
Db 11 CTTCCGCTCCGTCACCTGGTATTCTTAGTTCTCGAAAGACATTTTAGCTATGAA 70
|||
Qy 267 GluteuAlaArgAlaThrAsnGlyPheSerGluAlaAenLeuLeuGlyGlnGlyPhe 286
|||
Db 71 GAATTAGCGAGCAACGATGGCTTCGGAAGTTAACCTTCCTGGACAAGGTGGTTT 130
|||
Qy 287 GlyTyrrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys 306
|||
Db 131 GGGTACGTACACAAGGAGTTCTCCCTAATGGGAAGGAGTAGCAGTAAGCAACTCAAG 190
|||
Qy 307 ValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuLeuSerArgVal 326
|||
Db 191 GCTGGAAGTGGCAAGGCGAGAGAAATTCAGGCTGAAGTTGAGATCATTTAGCGCGTC 250
|||
Qy 327 HisHisArgHisLeuValSerLeuValGlyTyrrCysIleAlaGlyAlaLysArgLeuLeu 346
|||
Db 251 CATCACAACATCTCGTCTCATTCGTCGATCTGTTTCTGGACAAATGAATGCTT 310
|||
Qy 347 ValTyrrGluPheValProAsnAsnLeuLeuGluLeuHisLysGlyGluArgPro 366
|||
Db 311 GTTTATGAGTTGTTTCCAAACAACACCTTGGAGTTTCACTTGCATGGGAAGGCGACTG 370
|||
Qy 367 ThrMetGluTyrrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrr 386
|||
Db 371 ACCATGATTTGGCGACGAGGATGAAATTCCTTTAGGATCTGCANAAGGACTGGCATAT 430
|||
Qy 387 LeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeu 406
|||
Db 431 CTTTCATGAGATTGTCATCTTAAGATCATTCACCGTGATTAAGCGCGCTAATATCTG 490
|||
Qy 407 IleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAsp 426
|||
Db 491 TTGGATTTCAGATTGAAGCAAGGTTGCTGATTTGGACTAGCGAAAATTTGCTTCGGAT 550
|||
Qy 427 ThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrrLeuAlaProGluTyrr 446
|||
Db 551 GTCNAACGCGACGCTCCACCAGGCGTATGGGTACTTTTCGGGTATTTAGCCCTTGAGTAT 610
|||
Qy 447 AlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValVal 464
|||
Db 611 GCTTCAAGTGGAAAGCTCACTGATAATCAGATGTTTCTCTTCGGGGTCATG 664
|||
```

## RESULT 14

CA072174  
LOCUS CA072174 695 bp mRNA linear EST 23-SEP-2003  
DEFINITION SCCMAM1004A11.g AM1 Saccharum officinarum cDNA clone SCCMAM1004A11  
5' mRNA sequence.  
ACCESSION CA072174  
VERSION CA072174.1 GI:34924325  
KEYWORDS EST.  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Saccharum.  
REFERENCE 1 (bases 1 to 695)  
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

TITLE  
JOURNAL  
COMMENT

The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: Clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: 004 row: A column: 11  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers

FEATURES  
source

1..695  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCCMAM1004A11"  
/lab\_host="DH10B"  
/clone\_lib="AM1"

/note="Organ: Apical meristem and tissues surrounding of  
mature plants; Vector: pSport1; Site\_1: SalI; Site\_2:  
NotI; An unidirectional cDNA library generated from  
[Apical meristem and tissues surrounding of mature  
plants]. cDNA was prepared from polyA+ mRNA using  
SuperScript Plasmid System Kit (Invitrogen). The  
double-strand cDNAs were fractionated in a sepharose  
CL-2B 40cm-columns and fragments sizing between 0.8 and  
1.5 Kb were directionally cloned into the vector. Details  
of each source of RNA and library construction can be  
obtained at <http://sucest.lad.ic.unicamp.br/public>"

## ORIGIN

Alignment Scores:  
Pred. No.: 1.13e-23 Length: 695  
Score: 1013.00 Matches: 195  
Percent Similarity: 91.85% Conservative: 19  
Best Local Similarity: 83.69% Mismatches: 16  
Query Match: 29.34% Indels: 3  
DB: 13 Gaps: 1

US-10-086-464-2 (1-647) x CA072174 (1-695)

```
Qy 271 AlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrrValHis 290
|||
Db 4 GCAACCAATGGGTTTCTGACGCTAATCTGCTGGGCAAGCGGTTTGGGTTTGTTCAC 63
|||
Qy 291 LysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGly 310
|||
Db 64 AAGGAGTGTCTACCAATGTCACAGAGTTGCTTTAAGCAGTTAAGAGATGGAGTGC 123
|||
Qy 311 GlnGlyGluArgGluPheGlnAlaGluValGluLeuIleSerArgValHisArgHis 330
|||
Db 124 CAGGAGAACCGAGTTTCAGGCAGAGTTTGAGATTATCAGCAGATACATCAAGCAT 183
|||
Qy 331 LeuValSerLeuValGlyTyrrCysIleAlaGlyAlaLysArgLeuLeuValTyrrGluPhe 350
|||
Db 184 CTTGTATCTTGGTGGCTATTGCATTTCTGGAGCCCAAGGTTCGTCGTACAGATT 243
|||
Qy 351 ValProAsnAsnLeuLeuGluLeuHisGlyGluGlyArgProThrMetGluTyrr 370
|||
Db 244 GTTCCAAACAATCATTTGGAAATTCACATTAAGGAAAGGAGCAGACCTTGATGG 303
|||
Qy 371 SerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrrLeuHisGluAsp 390
|||
Db 304 CCGACAAGACTAAATTTGCTCTGGGTTCTGCGAAGGATTTGGCATATCTTCATGAAGAT 363
|||
Qy 391 CysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuLeuAspPheLys 410
|||
Db 364 TGCCATCTTAAGATCATTCATCGTGACATAAAGCGCTCAAAATATTCTTCTTGATCTAAGA 423
|||
```

Qy 411 PheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHis 430  
 Db 424 TTTGAAGCTAAGGTGGCAGATTTTGGACTTGCATAAATTCACCTTCTGTATACAAACACCCAT 483  
 Qy 431 ValSerThrArgValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGly 450  
 Db 484 GTTTCACACAGATAATGCGGCACATTTGGGTACCTAGCACCTGTAGTATGCTTCTTGGC 543  
 Qy 451 LysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThr 470  
 Db 544 AAGCTCACTGAGAAATCAGATGCTCTCTCTTTTGGAGTAAATGCTCTCTTGGAGCTAATAACT 603  
 Qy 471 GlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuValAlaSerTyrAla 490  
 Db 604 GCGCGCGCTCTCT-----GGTAATTCAGACAGGCAGATGACAACTTTGGGTGACTGGGCA 657  
 Qy 491 Arg-ProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502  
 Db 658 AGGCGCTTTGATGAATAAGCATTTGAGGATGGTAAT 694

## RESULT 15

BU099573

## LOCUS

DEFINITION WHE3309\_A09\_A172S Chinese Spring wheat drought stressed root cDNA library Triticum aestivum cDNA clone WHE3309\_A09\_A17, mRNA

## ACCESSION

BU099573

## VERSION

EST.

## KEYWORDS

Triticum aestivum (bread wheat)

## SOURCE

Triticum aestivum

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Pooidae; Triticeae; Triticum.

1 (bases 1 to 715)

Anderson, O.D., Chao, S., Close, T.J., Crossman, C., Fenton, R.D.,

Lazo, G.R., Nguyen H.T., Pham, J., Rausch, C.J., Turuspekov, Y.,

Wilson, C., Woo, J., and Zhang, D.

The structure and function of the expressed portion of the wheat

genomes - Chinese Spring drought stressed root cDNA library

Unpublished (2002)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderson@pw.usda.gov

Sequences have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq primer: SK primer.

Location/Qualifiers

1..715

/organism="Triticum aestivum"

/mol\_type="mRNA"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE3309\_A09\_A17"

/tissue\_type="root"

/dev\_stage="Full tillering stage"

/lab\_host="E. coli SOLR"

/clone\_lib="Chinese Spring wheat drought stressed root

cDNA library"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

Site 1: EcoRI; Site 2: XhoI; Plants were grown under

normal conditions. Then drought stressed to 80%, 70% and

60% RWC at Texas Tech University (D. Zhang in H.T. Nguyen

lab). Total RNA was prepared separately for roots

collected at the three different drought conditions. Equal

amounts of total RNA were pooled from all three samples,

poly(A) RNA were purified, one cDNA library was made, and

the cDNA clones were in vivo excised to give phluescript

SK(-) phagemids in the TJ Close lab at the University of

California, Riverside (Fenton, Turuspekov). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

## ORIGIN

Alignment Scores: 1.04e-22 Length: 715  
 Pred. No.: 983.00 Matches: 184  
 Score: 89.70% Conservative: 25  
 Percent Similarity: 78.97% Mismatches: 24  
 Best Local Similarity: 28.47% Indels: 0  
 Query Match: 13 Gaps: 0  
 Db: 0

US-10-086-464-2 (1-647) x BU099573 (1-715)

Qy 276 SerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuPro 295  
 Db 11 TCCGATGCTAATCTTCTCGGGCAAGGTGGCTTTGGATTGTTTTCACAAAGGAGTGTGCCA 70  
 Qy 296 SerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGlu 315  
 Db 71 GATGGCAGAGAAGTTGCTGTGAAGCAATTAAGAGATGGAAGTGGGCGAGAGCGGTGAG 130  
 Qy 316 PheGlnAlaGluValGluIleIleSerArgValHisArgHisLeuValSerLeuVal 335  
 Db 131 TTCAGGCAGAGGTTGAGATTATCAGCCGAGTATCATATAACATCTCGTGACATTTGGTT 190  
 Qy 336 GlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsn 355  
 Db 191 GGTATTGCAATTTCTGAAGACAAGAGGTGCTTGTCTATGAGTTTGTTCCTCAATAACACG 250  
 Qy 356 LeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuLys 375  
 Db 251 TTGATATTCCATATACATGGAAGCGGTGGACCAACTATGGAGTGGCTTCAAGACTACGT 310  
 Qy 376 IleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIle 395  
 Db 311 ATTGCTTTGGGTTCTCGAAGGGATTTGGGTATCTTCACGAAGACTGCGCATCCAAAGATC 370  
 Qy 396 IleHisArgAspIleLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysVal 415  
 Db 371 ATTCACTGTGACATAAAGGCATCAANTATCTTCTGGATTACAGATGTAAGCTAAGGTG 430  
 Qy 416 AlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsnThrHisValSerThrArgVal 435  
 Db 431 GCAGATTTTGGACTTGCAAAAGTTAACTCTGATAATAACACTCATGTGTTCCACCAGAGTA 490  
 Qy 436 MetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLys 455  
 Db 491 ATGGGCACATTTGGGTACCTTGCACAGAGTATGCTTCTCTGGCAAGCTAACTGAGAAA 550  
 Qy 456 SerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProVal 475  
 Db 551 TCAGATGCTTTTCTTTTGGAGTAATGCTTCTCGAGTTAATAACTGGGCGCGCTCTGTA 610  
 Qy 476 AspAlaAsnValTyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsn 495  
 Db 611 AGTTCAAAACAAGCGCATATGATGACAGCTTGGTTGCTGACTGGGCAAGCGCTTTGATGACA 670  
 Qy 496 ArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAsp 508  
 Db 671 CAAGCACTCGAGGATGTAATACAGATGCTTTTGTGGAT 709

Search completed: May 13, 2004, 01:33:29

Job time : 3622 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2004, 22:28:46 ; Search time 139 Seconds  
(without alignments)  
2583.117 Million cell updates/sec

Title: US-10-086-464-2  
Perfect score: 3453  
Sequence: 1 MSSAPSGTGTGPPSPNSNT.....REMEMGKIKRTGQYSGPSL 647

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues  
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10086464/runat\_11052004\_121815\_7866/app.query.fasta\_1.839  
-DB=issued Patents NA -QMT=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10086464 -CGN 1 1.85 @runat\_11052004\_121815\_7866 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	756	21.9	3239	4	US-09-228-986-9
2	744.5	21.6	1593	4	US-09-602-472A-1
3	740	21.4	2868	4	US-09-228-986-4
4	712	20.6	1488	4	US-09-579-182-4
5	703.5	20.4	2513	4	US-09-228-986-13
6	678	19.6	2432	4	US-09-228-986-7
7	667.5	19.3	1559	4	US-09-602-472A-3
8	604.5	17.5	4104	3	US-08-881-706-1
9	602	17.4	1224	4	US-09-579-182-3
10	600.5	17.4	3097	4	US-09-228-986-11
11	581.5	16.8	3590	1	US-08-587-889-1
12	581.5	16.8	3590	4	US-09-016-434-1093

13	581.5	16.8	3590	5	PCT-US96-09193-1	Sequence 1, Appli
14	578	16.7	2114	4	US-09-602-472A-5	Sequence 5, Appli
15	575	16.7	2749	1	US-08-265-628-1	Sequence 1, Appli
16	572	16.6	2571	1	US-07-717-331F-9	Sequence 9, Appli
17	572	16.6	2833	1	US-07-717-331F-1	Sequence 1, Appli
18	567	16.4	1554	2	US-08-587-680A-24	Sequence 24, Appli
19	566.5	16.4	2389	4	US-09-228-986-1	Sequence 1, Appli
20	562	16.3	2336	4	US-09-228-986-10	Sequence 10, Appli
21	559.5	16.2	2638	4	US-09-228-986-8	Sequence 8, Appli
22	546	15.8	2749	1	US-07-717-331F-4	Sequence 4, Appli
23	545	15.8	966	1	US-08-447-185-2	Sequence 2, Appli
24	545	15.8	2443	1	US-08-447-185-3	Sequence 3, Appli
25	543	15.7	2686	4	US-09-228-986-3	Sequence 3, Appli
26	531.5	15.4	5733	2	US-08-473-553A-1	Sequence 1, Appli
27	486	14.1	2943	4	US-09-503-922-2	Sequence 2, Appli
28	460.5	13.3	2649	4	US-09-228-986-12	Sequence 12, Appli
29	430	12.5	2568	4	US-09-228-986-2	Sequence 2, Appli
30	401	11.6	1926	4	US-09-249-585A-2	Sequence 2, Appli
31	401	11.6	1926	4	US-09-410-399-3	Sequence 3, Appli
32	401	11.6	2580	3	US-09-050-863-2	Sequence 2, Appli
33	401	11.6	2580	4	US-09-359-081-2	Sequence 2, Appli
34	401	11.6	5452	2	US-09-130-114-1	Sequence 1, Appli
35	401	11.6	8705	4	US-09-647-344A-14	Sequence 14, Appli
36	401	11.6	9600	3	US-08-910-647-1	Sequence 1, Appli
37	401	11.6	9600	4	US-09-620-925-1	Sequence 1, Appli
38	401	11.6	10596	1	US-07-884-811-15	Sequence 15, Appli
39	401	11.6	10596	1	US-07-885-971-15	Sequence 15, Appli
40	401	11.6	10596	1	US-08-087-783A-15	Sequence 15, Appli
41	401	11.6	10596	1	US-08-134-088B-15	Sequence 15, Appli
42	401	11.6	10596	2	US-08-194-087-15	Sequence 15, Appli
43	401	11.6	10596	5	PCT-US93-04648-15	Sequence 15, Appli
44	401	11.6	16080	4	US-09-724-566A-48	Sequence 48, Appli
45	395.5	11.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-228-986-9  
; Sequence 9, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Nieuwenhuizen, Timothy  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9  
; LENGTH: 3239  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
; US-09-228-986-9

Alignment Scores:  
Pred. No.: 1.88e-23  
Score: 756.00  
Percent Similarity: 51.93%  
Best Local Similarity: 34.99%  
Query Match: 21.89%  
DB: 4  
Matches: 3239  
Conservative: 190  
Mismatch: 92  
Indels: 132  
Gaps: 17

US-10-086-464-2 (1-647) x US-09-228-986-9 (1-3239)

Qy 89 ProProAlaProVal-ThrProThrArgAsnProProSerValProGlyProPr 108  
Db 1759 CTTCCCAACATTCGGCCCTTATTCTATCCCTCCCTATGCTTTT-----CC 1811  
Qy 108 oSerAsnProSerArgGluGlySerProArgProProSerProProSe 128

```
Db 1812 AGACAAAT-----GGAGGAACC----- 1827
Qy 128 rProSerSerAspGlyLeuSerThrGlyValValValGlyLeuAlaLeuGlyValAl 148
Db 1828 -----GCCATAAGCAAAAGGTGTGATAGTTGGGATCGCTATTGGCGGCACGGT 1874
Qy 148 aLeuLeuValLleValThrLeuLleCysLeuLeuCysLysLysLysArgArgAspG 168
Db 1875 TCTGGTCTTCGGCCTTGTTGTATTAGGGTTATATGCTATTCCACAAAAGAAACGGCGGA 1934
Qy 168 uLluAspAlaTyrValProProPro-----ProProProGlyProLysAl 184
Db 1935 GAAA---GCTCTCGAGTTGACACACACCTTCGCATCTCGGCACCCAGTGGGAAGATAG 1991
Qy 184 aGlyGlyProTyrGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSe 204
Db 1992 CGGAGGAGCGCCACACACTGAAAGGAGCAGCATGG----- 2025
Qy 204 rAspHisValValThrSerLeuProProProProLysAlaProSerProProArgGlnPr 224
Db 2025 ----- 2025
Qy 224 oProProProProProProPheMetSerSerSerGlyGlySerAspTyrSerAspAr 244
Db 2025 ----- 2025
Qy 244 gProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheTh 264
Db 2026 -----TTCTC 2030
Qy 264 rTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGln 284
Db 2031 CTATGATGAACCTTAAGAGGTGCACCAATAATTTCTCCGATAGCAATGAATTAGCTTCGG 2090
Qy 284 yGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysG 304
Db 2091 AGGATACGGAAGGTGTACAGGGAGGTCTCTCTGATGGTGCATATATAGCAATCAAAAG 2150
Qy 304 nLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluLleSe 324
Db 2151 AGCTCAGCAGCGGTTCGATCGAGGTGCAACCGAGTTCAAGACAGAAATCGAGCTGCTTC 2210
Qy 324 rArgValHisHisArgHisLeuValSerLeuValGlyTyrCysLleAlaGlyAlaLysAr 344
Db 2211 GCGGGTTTCATCACAAGAACTTGTGTGGCCTCATAGGATTCGTTCGAGCAAGAGAGCA 2270
Qy 344 gLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGlu 364
Db 2271 GATGTTGGTCTATGATATATATGCTTACGGGACGCTCAGGGATAGCTTGACAGAAATC 2330
Qy 364 yArgProThrMetGluTrpSerThrArgLeuLysLleAlaLeuGlySerAlaLysGlyLe 384
Db 2331 AGGCATTATCTTGATTGGAAGAGAGGCTTCGTATAGCTCTAGGTTTCGGCTAGGAGCT 2390
Qy 384 uSerTyrLeuHisGluAspCysAsnProLysLleHisArgAspLleLysAlaSerAs 404
Db 2391 AGCTTATCTGCACAACTCGCGAATCTCTCCAAATTTATCCACAGATGTCGAAGTCCACAA 2450
Qy 404 nLeuLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysLleAl 424
Db 2451 TATCTTGTGGACCAACATCTGACGGCCAAAGTTCGGGATTTTCGGTTTGTCCAACTGGT 2510
Qy 424 aSerAspThrAsnThr---HisValSerThrArgValMetGlyThrPheGlyTyrLeuAl 443
Db 2511 ATCGGACAGCGGGAAGGGGACGCTTCGACCAAGTGAAGGACGCTGGGCTATTTTGA 2570
Qy 443 aProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVa 463
Db 2571 TCCCGAATACATACATGATGATTCACAGCTGACAGAAAGCGCATGTGTACAGCTTCGGGGT 2630
Qy 463 lValLeuLeuGluLeuLleThrGlyArgArgProValAspAlaAsnValTyrValAs 483
Db 2631 GGTCAATGCTTGAGCTCATCTGCAAGCAACCACTTGAGNAGGCGCAAGTATGTCTGTCG 2690
```

```
Qy 483 pAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPh 503
Db 2691 CGAG-----ATTTCGACCCGCCATCGACAAGAACCAG-----GACTA 2729
Qy 503 eGluGly-----LeuAlaAspAlaLysMetAsnAsn---GlyTyrAspArgGluG 519
Db 2730 CTACGGCGTGGAGAAATGATGGACCCGTCCTCCTCAGGAGCATGGGCTAC---CTCGTCGG 2786
Qy 519 uMetAlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgPr 539
Db 2787 GTTCAGCAGGTTCTTGATTTGGGATTCGATGTCGAGGATCGGCTCGGACCGCC 2846
Qy 539 oArgMetSerGlnLleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnG 559
Db 2847 CACAATGAGCGAGGTGCTGAAGCGATCGAGACCATGTTGCAG-----AACGA 2894
Qy 559 uGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAs 579
Db 2895 TGGGATACACACCACCTCGACGTCGTCGCTCGTCG-----GCGACGGACTTTGG 2945
Qy 579 pSerSerGln-----TyrAsnGluAspMetLysLysPheArgLysMe 593
Db 2946 GTCGACGAGGGCGCTCTCTCGGCATCCGTACACAGCATCTTACCAG----- 2994
Qy 593 tAlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTy 613
Db 2995 -----AAGGAAGTTAGCTATAGCATTCCTTTGATTATAGT-----GGTGGATA 3038
Qy 613 rGlyLeu 615
Db 3039 TGGACTA 3045

RESULT 2
US-09-602-472A-1
; Sequence 1, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/140,876
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Helianthus annuus
; FEATURE:
US-09-602-472A-1

Alignment Scores:
Pred. No.: 2,9e-23 Length: 1593
Score: 744.50 Matches: 172
Percent Similarity: 57.07% Conservative: 66
Best Local Similarity: 41.25% Mismatches: 116
Query Match: 21.56% Indels: 63
DB: 4 Gaps: 9

US-10-086-464-2 (1-647) x US-09-602-472A-1 (1-1593)
Qy 235 SerSerGlyGlySerAspTyrSerAspArgProValLeuProProSerProGlyLeu 254
Db 181 TCGTCGGGTTTCGAAAGACGGGTCGACGGTCAGATT----- 216
Qy 255 ValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGly 274
Db 217 -----GCTGCTCATACATTCACTTCGCGGAGCTTGACGTGCAACAACAAAT 264
```



Qy	287	GlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLys	306
Db	1188	GGAAAGGTGTACAAAGGCGCCTTGCAGATGGTTCTTGGTGGCTGTAAACCGTCTGAAG	1247
Qy	307	---ValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIleSerArg	325
Db	1248	GAAGAGCGTACACCGGTTGAGAGATTCCGATCTTCAACACAGAAGTGGAGATGATAAGCATG	1307
Qy	326	ValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeu	345
Db	1308	GCAGTACATAGGAACCTCTTCGNCTACGTGGATCTTCGATGACACCCACTGAACGGCTG	1367
Qy	346	LeuValTyrGluPheValProAsnAsnAsnLeuGlu-----LeuHisLeuHisGlyGlu	363
Db	1368	CTTGTTTATCCCTACATGGCCAAATGGAAAGTGTGCTTCATGCCTACGACGAGAGGGCACAA	1427
Qy	364	GlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly	383
Db	1428	AATGACCCACCCCTTAGATTGGCCAACTCGCAAGCGCATAGCATTGGGTTCTGCAAGAGGT	1487
Qy	384	LeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSer	403
Db	1488	CTCTCTACTTGCATGATCATTTGTGATCCTAAGATATTACCACGGGATGTCAGGCTGCT	1547
Qy	404	AsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle	423
Db	1548	AACATCTTACTGGATGAAGATAATAGAGCAAGTGGTGGGGGATTTTGGCTTGGCAAACTT	1607
Qy	424	AlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla	443
Db	1608	ATGGATTATTAAGGACACACATGTTTACACCGCTGTTCTGCGAACCATTTGGCCCATAGCA	1667
Qy	444	ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyVal	463
Db	1668	CCTGAGTACCTTTTCTACTGGAAGTCTTCGSAAGACAGACAGCTATTTTGATTTGGAATC	1727
Qy	464	ValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnValTyrValAsp	483
Db	1728	ATGTTGCTGGAACCTATTATACGGGACAAACGGGCATTTTGACCTTGCACTTTAGCAAAATGAT	1787
Qy	484	AspSer-----LeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGlnGly	501
Db	1788	GATCATGTCTGTTGCTTGACTGGGTTAAGGCTACTA-----AAAGAGAGA	1835
Qy	502	AspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAla	521
Db	1836	AGGCTTGATATGCTAGTTGATCTCTGATCTTTAAGAACAAATTATGTTGAAGCAGAGGTGAA	1895
Qy	522	ArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMet	541
Db	1896	CAACTTATCAAGTTGCTATTACTTTGTACACAGGGGTCCAAATGGATGACCAAAAGATG	1955
Qy	542	SerGlnIleValArgAlaLeuGluGlyAsn-ValSerLeuSerAspLeuAsnGluGlyMe	561
Db	1956	TCTGAAGTGTAGGATGTTGGAAGGGGATCGCTTACTTGAGAGATGGGAGGAATGGCAA	2015
Qy	561	tArgProGlyGlnSerAsnValTyrSerSerTyrGly	573
Db	2016	AAAG-----TGAAAGTGTGACGGA	2034

## RESULT 4

```

RESUL 4
US-09-579-182-4
; Sequence 4, Application US/09579182
; Patent No. 6500628
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: MN1-161
; CURRENT APPLICATION NUMBER: US/09/579,182
; CURRENT FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 7

```

```
Qy 476 AspAlaAsnValTyrValAspSerLeuValAspTyrAlaAspProLeuLeuAsn 495
Db 808 GATAATGCTGAGCACCAGGAGCACAACCTTGTGCGATGGCTAGGCGGTGTC--- 864
Qy 496 ArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyr 515
Db 865 -----AAGATCGTAGAAGTTTCCGAAGATGGCGGATCCATCGCTCAAGGGCGGTAT 918
Qy 516 AspArgGluGluMetAlaAlaValAlaCysAlaAlaAlaCysValArgHisSerAla 535
Db 919 CCAATGCTGCTGCTATATCAAGCACTTGCAGTGTGCGCAATGTGTTTACAGAACAAGCA 978
Qy 536 ArgArgArgProArgMetSerGlnLeuValArgAlaLeuGluGlyAsnValSerLeuSer 555
Db 979 GCGACAAGACCACTGATTGGCGACGTGTGACAGCTCTTAACATCTTAGCTTCGCAACG 1038
Qy 556 ---AspLeuAsn-----GluGlyMetArgProGlyGlnSerAsnValTyr--- 569
Db 1039 TTTGACCAAAAGCCACCAAGCGGTCAAAACAGTAGAAGTGGGAGTGGGCCACCATTTATC 1098
Qy 570 -----SerSerTyrGlyGlySerThrAspTyrAsp----- 579
Db 1099 AGAACAGGATGATCGGAGAGCTTGGGAGATGGGAGTAGCTGGATAGCTTCGCAGAG 1158
Qy 580 -----SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMet 593
Db 1159 ACTCGAGTCGTTAGGTCACCGCACTCACAAGAACTCTCCTGATTACAGAAGAGG 1218
Qy 594 AlaLeuGlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSer----- 611
Db 1219 GATATG---GTCAGGGAAGTCAATGACAGATCAGAAGGTGGGAGCGAGCAGAGGCGGG 1275
Qy 612 -----AspTyrGlyLeuTyrProSerGlySerSerGlyGlnGlnThrArg 628
Db 1276 TCAGTAGAAATGGGGATTA-----AGCGATTGGGAAGGCAAGAATCACAG 1323
Qy 629 GluMetGluMetGlyLysIleLysArgThrGlyGlnGly 641
Db 1324 AGAGGAGCCCGCGGAGTGTGGGAGATCATCGAGAGGC 1362
```

## RESULT 5

```
US-09-228-986-13
; Sequence 13, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2513
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-13
```

```
Alignment Scores:
Pred. No.: 2,21e-21 Length: 2513
Score: 703.50 Matches: 178
Percent Similarity: 49.52% Conservative: 81
Best Local Similarity: 34.03% Mismatches: 179
Query Match: 20.37% Indels: 85
DB: 4 Gaps: 14
```

US-10-086-464-2 (1-647) x US-09-228-986-13 (1-2513)

```
Qy 84 ThrThrProGlySerProAlaProValThrProThrArgAsnProProSer 103
Db 795 ACTGTCCTGCTCCTCCTCCCAACCATATATCTCCACCT----- 833
```

```
Qy 104 ValProGlyProProSerAsnProSerArgGluGlyGlySerProArgProProSerSer 123
Db 834 -----CCACCTTCTTAC-----AATCGGCCACCCCAAGTCC 863
Qy 124 ProSer-----ProProSerProSerSerAspGly-----Leu 134
Db 864 TCAAAATGTTGGTCCCTTCATCAGTGGGTCAAGGGTGGAAACAGCAATAAGAAATCTCTG 923
Qy 135 SerThrGlyValValValGlyIleAlaIleGlyValAlaLeuLeuValIleValThr 154
Db 924 AGTGGTGGTGCATAGTGGGTATAATATTTCGACTTATTTCGACTGTGTGCTGCTATA 983
Qy 155 LeuIleCysLeuLeuCysLysLysLysLysArgArgArgGluGluAspAlaTyrTyrVal 174
Db 984 TTAGAGTTATTATTATGACAGTAACTCTCTTAGAAGAGAGCAGGAT----- 1031
Qy 175 ProProProProProGlyProLysAlaGlyGlyProTyrGlyGlnGlnGlnGln 194
Db 1032 -----GAAGAAAACTA 1043
Qy 195 TrpArgGlnGlnAsnAlaThrPro-----ProSerAspHisValValThrSerLeuPro 212
Db 1044 AGCAATCGTGTGCTTTCACCCCTTATCTCCCTCGATGCTGAATTAATTGAAAGAGAGT 1103
Qy 213 ProProProLysAlaProSerProArgGlnProProProProProProProPhe 232
Db 1104 CCAGAGCAAAAGTCACTCATCACTCTTGAATCGCTCTTAAGCCTCCCTCTTGAA 1163
Qy 233 MetSerSerSerGly-----GlySerAspTyrSerAspArgProValLeu 247
Db 1164 CGCAACAAGTCTACAGGGGACAAAGGCTTCGGAAGTATTTTTCAGTAGAAGAGCACTAA 1223
Qy 248 ProProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGlu 267
Db 1224 AACCCA-----ATATCAACAACCTGAATATTTTATTTCAGAC 1259
Qy 268 LeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyPheGly 287
Db 1260 CTGCAATATGGCAACAATAGTTTGTAGTCAAGATAACTTATTGCGGAGGTGCTCTTGA 1319
Qy 288 TyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysVal 307
Db 1320 CGAATCTACCGAGCAGAGTTCAGATGGAAGATTTTGGCAGTGAAGAAATTTGGACACT 1379
Qy 308 Gly-----SerGlyGlnGlyGluArgPheGlnAlaGluValGluIleIleSerArg 325
Db 1380 TCTACGCTGTCCCTACAAAGCCCTGAAGACTTCTCGACCGCAGTATCTAATATATCGCGC 1439
Qy 326 ValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeu 345
Db 1440 CTACATCATCTTAACATTACAGAACTAGTGGTTATTTCACAGAAACATGACCAATACCTT 1499
Qy 346 LeuValTyrGluPheValProAsnAsnAsnLeu-----GluLeuHisLeuHisGlyGlu 363
Db 1500 CTTGTGTATGAATATTTTCGACAATGGATCACTCTATCAGCTATTTCACATGCGAGATGAG 1559
Qy 364 GlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGly 383
Db 1560 ACTACTAGAAATTTGCTTGGAAACATTCGTGTAAAGATTTGCGTGGGTTCAGCTCGAGTT 1619
Qy 384 LeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSer 403
Db 1620 TTAGAGTATTTCGATGAAGTTTGTCTCCATCTATTGTGCAATAAAAAATTCAGTCTGCT 1679
Qy 404 AsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIle 423
Db 1680 AATATTTTGTGATGATGATTTCACACCTCGTCTGCTGCGAGCTGTGGAAATTCGCG- 1733
Qy 424 AlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443
Db 1734 GCCCTCAATCAAATCTGAGCGTCAGGTTTCAGGTGCTGGGTTCATTTCGGGTACAGTGCT 1793
```

Qy	444	ProGlnTyrAlaAlaSerGlyLeuThrGluLysSerAspValPheSerPheGlyVal	463
Db	1794	CCTGAATATGTCATGTCAGGAATCTATACATGAAGACGATGTTATAGTTTGGAGTG	1853
Qy	464	ValLeuLeuGluLeuIleThrGlyArgProValAspAlaAsnAsnValTyrValAsp	483
Db	1954	GTAATGCTTGAGCTTTTGACAGGCGGAGGCCCTTGGATAGTTCAGAAGACAGGTGCAGAA	1913
Qy	484	AspSerLeuValAspTirAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPhe	503
Db	1914	CAGTCATTAGTAAGATGGCCACCCCTCAGCTTCAT-----GATATT	1955
Qy	504	GluGlyLeuAla-----AspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet	520
Db	1956	GATGCATTGGCAAGATGGTTGATGCCAGCTTGAAGGGAAGTTATCTCTGCAGAAATCTCTC	2015
Qy	521	AlaArgMetValAlaAlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArg	540
Db	2016	TCGGCCTTTCGTGATATTATTGGCCCTCTGCATTGAGCCTGAACACAGAAATTCGCTCTCCA	2075
Qy	541	MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly	560
Db	2076	ATGTCGAAGTGGTGCAAGCATTCGTTGCTATGATGCAACGGGCTACGCTCCAAATAAGAGG	2135
Qy	561	MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer	580
Db	2136	ATGACAGAGNATGAA-----ATGCAGACCACGATCCT	2168
Qy	581	SerGlnTyr	583
Db	2169	GCAGATTAT	2177

## RESULT 6

```

US-09-228-986-7
; Sequence 7, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2432
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-228-986-7

```

US-10-086-464-2 (1-647) x US-09-228-986-7 (1-2432)

Qy	254	LeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaAtrGAlaThrAsn	273
		:::	
Db	1046	ATTGTCTAGGCAATTCGAAG--AGATTCTCATTCAAAGAGTTACGTGTTCACATAAT	1102
		:::	
Qy	274	GlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyVal	293
		:::	
Db	1103	AATTTTAGTACCAAGAAATATTTTAGGAGTAGGAGGATATGCAATGTCTATAAAGGATTC	1162
		:::	
Qy	294	LeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer---GlyGlnGly	312
		:::	
Db	1163	CTACAAGATGGCACTATATAGTAGCAATAAAGGTTGAAGATGCTAATGTGGGAGGAGGA	1222
		:::	

RESULT 7

```

RESULI 7
US-09-602-472A-3
; Sequence 3, Application US/09602472A
; Patent No. 6608240
; GENERAL INFORMATION:
; APPLICANT: Bidney, Dennis L.
; APPLICANT: Hu, Xu
; APPLICANT: Li, Guihua
; TITLE OF INVENTION: Sunflower Disease Resistance Genes
; FILE REFERENCE: 35718/200630
; CURRENT APPLICATION NUMBER: US/09/602,472A
; CURRENT FILING DATE: 2000-06-23

```





```
Qy 236 SerGlySerAspTyrSerAspArgProValLeuProProSerPro-----Gly 253
Db 364 AAAGACGGATCGAGGTATCTCTGCTAATCCTTTGACAGCTCCATCTCTCTGCTGTT 423
Qy 254 LeuVal-----LeuGlyPheSerIysSerThrPheThrTyrGluGluLeuAla 269
Db 424 CTTCTCTGAGTTTCTCACCTTGGATGGGA---CATTGGTTCACTCTTAGAGATCTTCAG 480
Qy 270 ArgAlaThrAenGlyPheSerGluAlaAenLeuGlyGlnGlyGlyPheGlyTyrVal 289
Db 481 ATGCTACTAATCAGTTTCAAGGATATATATCATCGGTGATGTTGGATATGGAGTTGTT 540
Qy 290 HisIysGlyValLeuProSerGlyIysGluValAlaValIysGlnLeuIysValIysSer 309
Db 541 TACCGCGTAACCTTGTAAATGGTACTCTCTGCTGTTAAAGATTGCTCAACAATTTA 600
Qy 310 GlyGlnGlyGluArgGluPheGlnAlaGluValGluLeuIleSerArgValHisIleArg 329
Db 601 GGACAAGCTGATAAAGACTTCAGAGTTGAAGTTGAAGCTATAGGTCAAGTTTCGACATAAA 660
Qy 330 HisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaIysArgLeuLeuValTyrGlu 349
Db 661 AACTTGGTCCCTCTTCGGATATATGATGAAGAACGACGAGG-----705
Qy 350 PheValProAenAenAenLeuGluLeuHisGlyGlyGlyArgProThrMetGlu 369
Db 705 -----705
Qy 370 TrpSerThrArgLeuIysIleAlaLeuGlySerAlaIysGlyLeuSerTyrLeuHisGlu 389
Db 706 -----CTCGCGTACTCTTCAGGAG 723
Qy 390 AspCysAenProIysIleIleHisArgAspIleIysAlaSerAenIleLeuIleAspPhe 409
Db 724 GCGATTGAGCAAAAGTGTGTCAGAGACATTAACTAGTAACTTCTGATTGATGAC 783
Qy 410 LysPheGluAlaIysValAlaAspPheGlyLeuAlaIysIleAlaSerAspThrAsnThr 429
Db 784 AANTCAATTCTAANATTTCTGACTTTTGGACTTGTCTAACTTCTTGGTGTGATAAGAGT 843
Qy 430 HisValSerThrArgValMetGlyThrPheGlyTyrIleAlaProGluTyrAlaAlaSer 449
Db 844 TTTATAACTACTAGAGTTATGGTACCTTCGGTTACGTAGCTCCAGAGATATCGAATTCC 903
Qy 450 GlyIysLeuThrGluIysSerAspValPheSerPheGlyValValLeuLeuGluLeuIle 469
Db 904 GGTCTTCTGAATGAGAAAGCGATGCTACAGCTTCGGGGTTGTACTCTTGGAAAGCTATA 963
Qy 470 ThrGlyArgArgProValAspAlaAenAenValTyrValAspAspSerLeuValAspTyr 489
Db 964 ACTGTTAGATATCGGTAGACTATGCTCGTCCACCACCGAGGTACATTTGGTGGAGTGG 1023
Qy 490 AlaArgProLeuLeuAenArgAlaSerGluGlnGlyAspPheGlyLeuAlaAspAla 509
Db 1024 CTGAAGATGATGTC-----CAACAAAGACGATCAGAAGAGTGGTTGATCCA 1071
Qy 510 LysMetAenAenGlyTyrAspArgGluGluMetAlaArgMetValAlaCysAlaAlaAla 529
Db 1072 AACCTTGAACAAACCATTCTACAAGTCTTTGAAAAGAACACATTATGACTGCTTTCAGA 1131
Qy 530 CysValArgHisSerAlaArgArgProArgMetSerGlnIleValArgAlaLeuGlu 549
Db 1132 TGTGTTGATCCATGTCAGAAAAGACCGAGGATGAGCCAGTTGACGATGCTCTTGAA 1191
```

## RESULT 10

```
US-09-228-986-11
; Sequence 11, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
```

```
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 3097
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-228-986-11
Alignment Scores:
Pred. No.: 4,95e-17 Length: 3097
Score: 600.50 Matches: 173
Percent Similarity: 46.08% Conservative: 80
Best Local Similarity: 31.51% Mismatches: 172
Query Match: 17.39% Indels: 126
DB: 4 Gaps: 13
```

```
US-10-086-464-2 (1-647) x US-09-228-986-11 (1-3097)
Qy 66 ProSerThrProGlySerProProPro-----LeuPro 76
Db 1523 CTTGCAAGTTCTGGACGCTCCAAACACAGCTTAGACGGAGTTTCTCCGACGGCTTGGG 1582
Qy 77 GlnProSerProAlaProThrThrProGlySerProProAlaProValThrPro--- 95
Db 1583 AGAGCTAAAGACCTTCACCTACTGGACTTGGAGATAAAGCTGTACAAAG-GTACCTTAC 1641
Qy 96 -----ProThrArgAenProProProSerValProGlyProPro 108
Db 1642 CAGATCTCTTGAACAGACAGAGCTTGGAGGTCAGAACTCAGGAACTTGTGTCTTCT 1701
Qy 109 SerAsnProSerArgGluGlySerProArgProProSerProSerPro----- 126
Db 1702 TCTCCACAACCGCGTGC-GGTGATGCATCTCTTCAATTTAGGACCCCAAGTT 1760
Qy 127 -----ProSerProSerSerAspGlyLeuSerThrGlyValValValGlyIleAla 143
Db 1761 ACATAGTTCTCCGAGAGAAACAAAGGGGACATAATCGTTAGCATATATCTCGAGCA 1820
Qy 144 IleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLeuIys 163
Db 1821 GTCGAGGAGTATCAGCTAGCTATTTTACTATCCGCTCTTCGTATTTCATGTACAGAAG 1880
Qy 164 ArgArgAspGluGluAspAlaTyrTyrValProProProProProGlyProIys 183
Db 1881 AGAGGAAGAACTGAA-----1895
Qy 184 AlaGlyGlyProTyrGlyGlyGlnGlnGlnTrpArgGlnGlnAenAlaThrProPro 203
Db 1895 -----1895
Qy 204 SerAspHisValValThrSerLeuProProProIysAlaProSerProProArgGln 223
Db 1895 -----1895
Qy 224 ProProProProProPheMetSerSerSerGlySerAspTyrSerAsp 243
Db 1896 -----ATGTCATATACGAA 1910
Qy 244 ArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerIysSerThrPhe 263
Db 1911 AGGCAGTCCGACAGCTGAGAAACTGGAAT-----GCAGCTAAGATTTT 1955
Qy 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAenLeuLeuGlyGln 283
Db 1956 TCCTACAAAGAGATCAAAACAGCTACAAACAACTTTAAAGAA-----GTCAATGGTCA 2009
Qy 284 GlyGlyPheGlyTyrValHisIysGlyValLeuProSerGlyIysGluValAlaValIys 303
Db 2010 GGAAGTTTGGATCCGCTGATCTAGTGGGAAACCTTCCAGTTGGAAACCTAGTTGCTGTGAA 2069
```

```
Qy 304 GlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIle 323
Db 2070 GTGCGGTTTGATAAAACCAACTGGTGCAGATTCTTTTCATAAATGAGGTTGCTCTCTTA 2129
Qy 324 SerArgValHisLeuArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLys 343
Db 2130 TCACAGTCCCCATCAGAACCTGTGCACTCTGGAGGATTTGTATGAGTCGACGGT 2189
Qy 344 ArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHisGlyGlu 363
Db 2190 CAGATTTTGTCTATGAATATCTACCGGTTGATCACTGGCTGATCAACTGTATGTTCCA 2249
Qy 364 Gly-----ArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySerAla 381
Db 2250 AACAGTAGGAAATCTCTCAAGTGGTTCGTAGACTCAAGATTGCTGTGAGTGTGCA 2309
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLys 401
Db 2310 AAAGACTGGACTATCTACATTAATGAAGCAATCTCGAATCATACACCGAGACATCAAG 2369
Qy 402 AlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 2370 TGCAGTAATATCTTGGACAGAGATGAATGCAAGACTTTGCGACTTTGGGCTCTCT 2429
Qy 422 Lys-----IleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPhe 439
Db 2430 AAGCAAAATGATCCAGCCAGAGCGCA---ACTCACGTGACCACTGCTCAAGGGGACACAGCT 2486
Qy 440 GlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPhe 459
Db 2487 GGTACCTCGACCTGATATATCTACCCACCAACACTTACAGAGAAAGCGAGCTCTAT 2546
Qy 460 SerPheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsn 479
Db 2547 AGCTTTGAGTGTGCTTTTGGAGCTCATCTGTGACGAGAGCCGTTAAATCAATTCAGGA 2606
Qy 480 ValTyrValAspAspSerLeuValAspTrpAlaArgProLeuLeuLeuAsnArgAlaSerGlu 499
Db 2607 ACTCCAGATTCCTTCAAATTTGGTTTATGGCAAGCCCTACTTG-----CAG 2654
Qy 500 GlnGlyAspPheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGlu 519
Db 2655 CGAGGTGCATTGAG---ATAGTGATGAGAGTTAGGGGAAGTTTCGATGTGGAAGC 2711
Qy 520 MetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgPro 539
Db 2712 ATGAGAAAGTGGCAAAATCGCTGTGAGTCTGTAGAGAGGATGCATCACTAAGGCCA 2771
Qy 540 ArgMetSerGlnIleValArgAlaLeuGluGly-----AsnValSerLeuSerAspLeu 557
Db 2772 ACCATTGCACAGATACTCTGTGCTCAAGAGGCTTACAGCAATTCAGCTCTCT----- 2825
Qy 558 AsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAsp 577
Db 2826 -----TATCTTGAGGCTCTGGACATGTGAAC 2852
Qy 578 TyrAspSerSerGlnTyrAsnGluAsp 586
Db 2853 TGAATCTACTCGTATTACCAACCAAGAC 2879
```

## RESULT 11

US-08-587-889-1  
; Sequence 1, Application US/08587889  
; Patent No. 5654397

## GENERAL INFORMATION:

; APPLICANT: CAO, Zhaoan  
; APPLICANT: CROSTON, Glenn E.  
; APPLICANT: GOEDEL, David V.  
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED  
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/587,889  
; FILING DATE:

## CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman Ph.D., Richard Aron  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-60916  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 494-8700  
; TELEFAX: (415) 494-8771

; INFORMATION FOR SEQ ID NO: 1:

## SEQUENCE CHARACTERISTICS:

; LENGTH: 3590 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-587-889-1

## Alignment Scores:

Pred. No.: 3,48e-16 Length: 3590  
Score: 581.50 Matches: 208  
Percent Similarity: 38.68% Conservative: 86  
Best Local Similarity: 27.37% Mismatches: 239  
Query Match: 16.84% Indels: 229  
DB: 1 Gaps: 28

US-10-086-464-2 (1-647) x US-08-587-889-1 (1-3590)

```
Qy 1 MetSerAlaProSerPro-GlyThrGlySerProSerProSerProSerAsnSerTh 20
Db 80 ATGCCGGGGGCGCGCCCGGGGAGCCCGCAGACCCCGCGCGCCAGCATTTCTGTAC 139
Qy 20 rThrThrThrProProAlaSerAla-----ProProPr 32
Db 140 GAGGTGCCCGCTGGGTGATGTGCCGCTTCTACAAAGTAGAGCGCCCTGGAGCCCGCC 199
Qy 32 oThrThrProSerSerProPro----- 39
Db 200 GACTGTGCCAGTTTCGCCGCCCTGATCGTGCAGCAGCAGCCGAGCTGCGGAG 259
Qy 40 ----ProProSerThrIleProThrSerProProProSerSerArgSerThrProSerAl 58
Db 260 CGCTCCGGGAGCCAGCCAGCGCGCTCTGTGGCCCTGGATCAACCGCAACCGCGGTG 319
Qy 58 aProProProSerProPro-----ThrProSerTh 68
Db 320 GCCGACCTCGTGACATCTCTCAAGCAGCTGCTCGTGCAGGACATCATCACA 379
Qy 68 rProGlySerProProProLeuProGlnProSerProProAlaProThrThrProGlySe 88
Db 380 GCCTGGCACCCTCCCGCCCGCTTCGCTCCCAAGCAGCAGCTGCCCCGA-----GG 430
Qy 88 rProProAlaProValThrProThrArgAsnProProProSerValProGlyProPr 108
Db 431 CCCAGAGCATCCCTGCACCGCGG-----AGGCCGAGGCTGGAGCCCCCGGA----- 479
Qy 108 oSerAsnProSerArgGluGlyGlySerProArgProProSerProSerProProSe 128
Db 480 -----AGTTGCCATCCTCCAGCTCCACCTTCCTC-TCCCCAGCTTT 519
```

Qy 128 rProSerSerAspGlyLeuSerThrGlyValValValValGlyLeuAlaLeuGlyValAla 148  
Db 520 TCCAGGCTCCAGACCCATTCA----- 541  
Qy 148 aLeuLeuValLeuValThrLeuLeuLeuCysLeuLeuLeuValValValValValVal 168  
Db 541 ----- 541  
Qy 168 uGluAspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyProTy 188  
Db 542 -----GGGCTGAGCTGGC----- 556  
Qy 188 rGlyGlyGlnGlnGlnTrpArgGlnAsnAlaThrProProSerAspHisValVa 208  
Db 556 ----- 556  
Qy 208 lThrSerLeuProProProProProProProSerProProArgGlnProProProProPr 228  
Db 557 -----CTGGTTCCAAAGCCCTGCTTCCCTGTGGCTCCAGCCGCC 594  
Qy 228 oProProProPheMetSerSerGlySerAspTyrSerAspTyrProValLeuPr 248  
Db 595 ATCTCCAGCCCTTCTTACCAAGCCAGCC-----CC 627  
Qy 248 oProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGlu----- 266  
Db 628 AGAGAGCTCAGTCTCCCTCTGAGGAGCCGCCCTCTCCGTTTGTGCGCCCTCTG 687  
Qy 267 -GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyph 286  
Db 688 TGAGATTTCCTGGGGGACCCCAACTTCTCGAGGAGCTCAAGATCGGGAGGGTGGCTT 747  
Qy 286 eGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuL 306  
Db 748 TGGTGTGGTGTACCGGGCGGTGATG---AGGAACACGGTGTATCTGTGAAGAGCTGAA 804  
Qy 306 sValGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGluIl 322  
Db 805 GGAGAACGCTGACCTGGAGTGAGTGCAGTGAAGCAGAGCTTCTGACCGAGGTGGAGCA 864  
Qy 322 eLysSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysLeuAlaGlyAl 342  
Db 865 GCTGTCCAGGTTTCGTACCCCAACATTTGTGGACTTTGCTGGCTACTGTCTCAGAACGG 924  
Qy 342 aLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGlu-----LeuHisLe 360  
Db 925 CTCTACTGCTGGTGTACGGCTTCTCTGCCACACGGCTCCCTGGAGACCGCTCTCCACTG 984  
Qy 360 uHisGlyGlyArgProThrMetGluTrpSerThrArgLeuLysLeuAlaLeuGlySe 380  
Db 985 CCAGACCCAGCCCTGCGCCACCTCTCTCTGCGCTCAGCGACTGGACATCTCTTCTGGGTAC 1044  
Qy 380 rAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysLeuLeuHisArgAspIl 400  
Db 1045 AGCCCGGGCAATTTCAGTTTCTACATCAGGAC---AGCCCCAGCCCTCATCCATGAGACAT 1101  
Qy 400 eLysAlaSerAsnLeuLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLe 420  
Db 1102 CAAGATTTCACAGCTTCTTGGATGAGAGCTGACACCCCAAGCTGGGAGACTTTGGCT 1161  
Qy 420 uAlaLysLeu-----AlaSerAspThrAsnThrHisValSerThrArg-- 434  
Db 1162 GGCCCGGTTACCGCTTTTGGCGGTCCAGCCCGCCAGCAGCAGTGTGGCGCCGAC 1221  
Qy 435 -----ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLe 452  
Db 1222 ACAGACAGTGGGGGACCTCTGCTGCTACCTCCCGAGGAGTACATCAAGACGGGAGGCT 1281  
Qy 452 uThrGluLysSerAspValPheSerPheGlyValValLeuLeuLeuLeuThrGlyAr 472  
Db 1282 GGCTGTGACACGACACCTTTCAGCTTTGGGGTGGTGTAGTGTAGAGACCTTTGGCTGTCA 1341  
Qy 472 gArgProVal-----AspAlaAsnAsnValTyrValAspAspSerLeuValAspTr 489

Db 1342 GAGGCTGTGAAGACGACCGTGTCCAGGACCAAGTATCTGAAAGAC---CTGTGGAA-- 1396  
Qy 489 pAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509  
Db 1397 -----GAGGAGGCTGAGGAGGCTGGAGTGGCTTTTGAGAGCACCCAGAG 1440  
Qy 509 aLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArgMetValAla----- 525  
Db 1441 CACACTGCAAGCAGGCTGTGCTGCAGATGCTGGGCTGCTCCCATCGCCATGCAGATCTA 1500  
Qy 525 ----- 525  
Db 1501 CAAGAAGCACCTGACCCAGCGCCGCGCTGTGCCACCTGAGCTGGGCGCTGGCGCTGGG 1560  
Qy 526 ----CysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnIl 544  
Db 1561 CCAGCTGCGCTGTGCTGCTGCACCCCGCGGCGCAAGAGGAGGCTCTCTATGACCCAGGT 1620  
Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564  
Db 1621 GTACGAGAGGCTAGAG-----AAGCTGCAGGCAGTGTGTGCGGGGTG---CCCGG 1668  
Qy 564 Y-----GlnSerAsnValTyrSerSer-- 571  
Db 1669 GCATTTGGAGCGCCAGCTGCATCCCGCTTCCCGCAGGAGAACTCTACGTGTCCAG 1728  
Qy 572 -----TyrGlySerThrAspTyrAsp----- 579  
Db 1729 CACTGGCAGACCCAGCTGGGCTGTCTCATGACGCCCTTGGCAGCGCCATCAGGAGC 1788  
Qy 580 -SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGl 599  
Db 1789 CAGTGGCCAGCAGCAGCAGCTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1848  
Qy 599 uTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer-- 618  
Db 1849 G-----AGCTAGCGCGCTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1902  
Qy 619 -----GlySerSerSerGluGlyGlnThrThrArgGl 629  
Db 1903 CCCTTGTGACCCAGACCCCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1961  
Qy 629 uMetGluMetGly-----LysLeuLysArgThrGlyGlnGlyTyrSerGlyPro 645  
Db 1962 AATCGAGCTGGGAGGAGTGGCCCGGAGGATCCCGGCGCCACAGCGCTGGAAGGAGTGGCCCT 2019

## RESULT 12

US-09-016-434-1093  
Sequence 1093, Application US/09016434  
Patent No. 6500938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Sellhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA: US/09/016,434  
FILING DATE: HEREWITH  
CLASSIFICATION:

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1093:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1220312
US-09-016-434-1093

Alignment Scores:
Pred. No.: 3,48e-16 Length: 3590
Score: 581.50 Matches: 208
Percent Similarity: 38.68% Conservative: 86
Best Local Similarity: 27.37% Mismatches: 239
Query Match: 16.84% Indels: 229
Dbs: 4 Gaps: 28

US-10-086-464-2 (1-647) x US-09-016-434-1093 (1-3590)

QY 1 MetSerSerAlaProSerPro-GlyThrGlySerProProSerProProSerAsnSerTh 20
Db 80 ATGGCCGGGGCGGGCGGGCGGGGAGCCCGCAGCGCCCGCGCGCCAGCACTTCTGTGAC 139
QY 20 rThrThrThrProProProAlaSerAla-----ProProPr 32
Db 140 GAGGTGCCGCCCTGGGTGATGTCGCGCTTCTACAAAGTGATGGAGCGCCCTGGAGCGCCG 199
QY 32 oThrThrProSerSerProPro----- 39
Db 200 GACTGGTGCCAGTTCGCGCCCTGATGTCGGCGACCGAGCCGAGCTGCGGCTGTCGAG 259
QY 40 -----ProProSerThrIleProThrSerProProProSerSerArgSerThrProSerAl 58
Db 260 CGCTCCGGGCGAGCGCAGCGCCAGCGTCTGTGGCCCTGGATCAACCGCAACGCCCGTGTG 319
QY 58 aProProProSerProPro-----ThrProSerTh 68
Db 320 GCCGACCTCGTGACATCCTCAGCACCCTGCAGCTGCTCCGTGCGCGGAGCATCATACA 379
QY 68 rProProSerProProProLeuProGlnProSerProProAlaProThrThrProGlySe 88
Db 380 GCCTGGCACCTCCCGCCCGCTTCGTCGCCGAGCACACTGCCCCCA-----GG 430
QY 88 rProProAlaProValThrProProThrArgAsnProProProSerValProGlyProPr 108
Db 431 CCCAGCAGCATCCCTGCACCGCCG-----AGGCCGAGGCTGGAGCCCGCGA----- 479
QY 108 oSerAsnProSerArgGluGlyGlySerProArgProProSerSerProProProSe 128
Db 480 -----AGTTGCCATCCTCAGCTCCACCTTCCTC-TCGCCAGCTTT 519
QY 128 rProSerSerAspGlyLeuSerThrGlyValValGlyIleAlaIleGlyValAl 148
Db 520 TCCAGCTCCAGACCCATTCA----- 541
QY 148 aLeuLeuValIleValThrLeuIleCysLeuLeuCysLeuLeuLysLysLysArgArgAspGl 168
Db 541 ----- 541
QY 168 uGluAspAlaTyrtYrValProProProProGlyProGlyProGlyAlaGlyGlyProTy 188

```

```

Db 542 -----GGGCTGAGCTCGGC----- 556
QY 188 rGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValVa 208
Db 556 ----- 556
QY 208 lThrSerLeuProProProProlyAlaProSerProProArgGlnProProProProPr 228
Db 557 -----CTGGTTCCAAGCCCTGCTTCCCTGTGGCTCCACCGCC 594
QY 228 oProProPheMetSerSerSerGlyGlySerAspTy-SerAspArgProValLeuPr 248
Db 595 ATCTCCAGCCCTTCTTCTACCAAGCCAGCC-----CC 627
QY 248 oProProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyGlu----- 266
Db 628 AGAGAGCTCAGTGTCTCTCCAGGAGCGCCGCTCTCCGTTTGTCTGCCCCCTCTG 687
QY 267 -GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyph 286
Db 688 TGAGATTTCGCGGGCACCCCAACTTCTCGAGAGAGCTCAAGATCGGGAGGGTGGCTT 747
QY 286 eGlyTy-ValHisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLy 306
Db 748 TGGGTGCGTGTACCGGGCGGTGATG---AGGAACACCGTGTATGCTGTGAAGAGCGTGAA 804
QY 306 svalGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGluI 322
Db 805 GGAGAACGCTGACCTGGAGTGGACTGCAGTGAAGCAGAGACTTCTCGACCGAGGTGGAGCA 864
QY 322 eIleSerArgValHisHisArgHisLeuValSerLeuValGlyTy-CysIleAlaGlyAl 342
Db 865 GCTGTCCAGGTTTCGTACCCAAACATGTGTGAGACTTGTCTGGCTACTGTCTCAGAACGG 924
QY 342 aLysArgLeuLeuValTyGluPheValProAsnAsnAsnLeuGlu-----LeuHisLe 360
Db 925 CTTCTACTGCTGTGTGTACGGCTTCCTGCCCAAGGGTCCCTCGGAGGACCCTTCCCATG 984
QY 360 uHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySe 380
Db 985 CCAGACCCAGCCCTGCCACCTCTCTCTGCTCAGCGACTCGACATCTCTTCTGGGTAC 1044
QY 380 rAlaLysGlyLeuSerTyLeuHisGluAspCysAsnProLysIleIleHisArgAspIl 400
Db 1045 AGCCCGGGCAATTTCAGTTTCTACATCAGGAC---AGCCCCAGCCTCATCTCCATGGAGACAT 1101
QY 400 eLysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLe 420
Db 1102 CAAGATTCCAACTCTCTCTGGATGAGAGCTGACACCCCAAGCTGGGAGACTTTGGCT 1161
QY 420 uAlaLysIle-----AlaSerAspThrAsnThrHisValSerThrArg-- 434
Db 1162 GCGCCGGTTACGCGCTTTTGGCGGTTCAGCCCGCAGCGACGAGCAGTGTGGCCGAG 1221
QY 435 -----ValMetGlyThrPheGlyTyLeuAlaProGluTyrAlaAlaSerGlyLysLe 452
Db 1222 ACAGACAGTGGGGGACCCCTGGCTTACCTCCCGAGGAGTACATCAAGACGGGAAGCT 1281
QY 452 uThrGluLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyAr 472
Db 1282 GCCTGTGGACGAGCACCTTCAGCTTTGGGGTGGTAGTGTAGAGACTTGGCTGGTCA 1341
QY 472 gArgProVal-----AspAlaAsnAsnValTy-ValAspSerSerLeuValAspTr 489
Db 1342 GAGGCTGTGAAGACGCGCGTGCAGGACCAAGTATCTGAAGAC---CTGGTGGAA-- 1396
QY 489 pAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509
Db 1397 -----GAGGAGCTGAGGAGTGGAGTGGCTTTGTGAAGACCCACGAG 1440
QY 509 aLysMetAsnAsnGlyTy-AspArgGluGluMetAlaArgMetValAla----- 525

```

```

Db 1441 CACACTGCAAGCAGGTCTGGTGCAGATGCTTGGGCTGCTCCATCGCCATGCAGATCTA 1500
Qy 525 -----
Db 1501 CAAGAAGCAGCTGGAGCCCGGCGCCCTCCACCTGAGCTGGGCGCTGGGCGCTGG 1560
Qy 526 ---CysAlaAlaAlaCysValArgHisSerAlaArgArgProArgMetSerGlnI 544
Db 1561 CCAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProG 564
Db 1621 GTACGAGAGGCTAGAG-----AGCTGAGGAGCTGTGTGGGGGGTG---CCCGG 1668
Qy 564 Y-----GlnSerAsnValTySerSer-- 571
Db 1669 GCATTGGAGGCGGCAGCTGCATCCCTCCCGGAGGAGAACTCTAGTGTCCAG 1728
Qy 572 -----TyGlyGlySerThrAspTyAsp----- 579
Db 1729 CACTGGCAGAGCCACAGTGGGGTGTCTCCATGGCAGCCCTGGCAGCGCCATCAGGAGC 1788
Qy 580 -SerSerGlnTyAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnG 599
Db 1789 CAGTCCCGCAGCAGCAGCAGCAGCTGCAGAGAGGCCCAACAGCCGCTGGAGAGTGACGA 1848
Qy 599 uTyAsnAlaThrGlyGluTySerSerAsnProThrSerAspTyGlyLeuTyProSer-- 618
Db 1849 G-----AGCTAGGCGGCTCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1902
Qy 619 -----GlySerSerSerGluGlyGlnThrThrArgG 629
Db 1903 CCCTGTGACCCAGCACCCTCAGGAGGCGGCTGTCTCAGGGGAGACAGGC-AGGAG 1961
Qy 629 uMetGluMetGly-----LysIleLysArgThrGlyGlnGlyTySerGlyPro 645
Db 1962 AATCAGCTGGGGAGTGGCCAGGATCCCGGCCACAGCCGCTGGAGGAGTGGCCCT 2019

RESULT 13
PCT-US96-09193-1
; Sequence 1, Application PC/TUS9609193
; GENERAL INFORMATION:
; APPLICANT: Tularik, Inc.
; TITLE OF INVENTION: INTERLEUKIN-1 RECEPTOR-ASSOCIATED
; TITLE OF INVENTION: PROTEIN KINASE AND BINDING ASSAY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09193
; FILING DATE: JUNE 5 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Serial No. 08/587,889
; FILING DATE: JAN 16 1996
; CLASSIFICATION:
; APPLICATION NUMBER: U.S. Serial No. 08/494,006
; FILING DATE: JUNE 23 1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: David J. Brezner
; REGISTRATION NUMBER: 24,774
; REFERENCE/DOCKET NUMBER: FP-62191-1

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US96-09193-1

Alignment Scores:
Pred. No.: 3,48e-16 Length: 3590
Score: 581.50 Matches: 208
Percent Similarity: 38.68% Conservative: 86
Best Local Similarity: 27.37% Mismatches: 239
Query Match: 16.84% Indels: 229
DB: 5 Gaps: 28

US-10-086-464-2 (1-647) x PCT-US96-09193-1 (1-3590)

Qy 1 MetSerSerAlaProSerPro-GlyThrGlySerProSerProSerProSerProSerTh 20
Db 80 ATGCCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 139
Qy 20 rThrThrThrProProAlaSerAla-----ProProPr 32
Db 140 GAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 199
Qy 32 oThrThrProSerProPro----- 39
Db 200 GACTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 259
Qy 40 ----ProProSerThrThrProProSerProProProSerProProSerProSerAl 58
Db 260 CGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 319
Qy 58 aProProProSerProPro-----ThrProSerTh 68
Db 320 GCCGACCTCGTGCACTCTCAGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 379
Qy 68 rProGlySerProProLeuProGlnProSerProProAlaProThrThrProGlySe 88
Db 380 GCTTGGCACCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 430
Qy 88 rProProAlaProValThrProProThrProProProProProProProProProPro 108
Db 431 CCCAGCAGCATCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 479
Qy 108 oSerAsnProSerArgGlyGlySerProProProProProProProProProProPro 128
Db 480 -----AGTTGCCATCTCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 519
Qy 128 rProSerSerAspGlyLeuSerThrGlyValValValGlyLeuAlaIleGlyGlyVal 148
Db 520 TCCAGGCTCCCGAGACCCATTCA----- 541
Qy 148 aLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysArgArgArgAspG 168
Db 541 ----- 541
Qy 168 uGluAspAlaTyTyValProProProProProProProProProProProProProTy 188
Db 542 -----GGGCGCTGAGCTCGCG----- 556
Qy 188 rGlyGlyGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisVal 208
Db 556 ----- 556
Qy 208 lThrSerLeuProProProLysAlaProSerProProProProProProProProPro 228

```

Db 557 -----CTGGTTCCAGCCCTCTTCCCTGTGGCTCCACCGCC 594  
Qy 228 oProProPhoMetSerSerSerGlyGlySerAspTyrSerAspArgProValLeuPr 248  
Db 595 ATCTCCAGCCCTTCTTACCAAGCCAGGC-----CC 627  
Qy 248 oProProSerProGlyLeuValLeuGlyPheSerIysSerThrPheThrTyrGlu----- 266  
Db 628 AGAGAGCTCAGTTCCTCTGCGAGGAGCCGCCCTCTCCGTTTGTCTGGCCCTCTGT 687  
Qy 267 -GluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPh 286  
Db 688 TGAGATTTCCTGGGACCCACACTTCTCGAGGAGCTCAAGATCGGGAGGTGGCTT 747  
Qy 286 eGlyTyrValHisIysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLy 306  
Db 748 TGGGTGCTGTACCGGGCGGTGATG---AGGAACACGGTGTATGCTGTGAAGAGCTGAA 804  
Qy 306 sValGlySerGly-----GlnGlyGluArgGluPheGlnAlaGluValGlu1 322  
Db 805 GGAGAACCTGACCTGGAGTGCAGTGAAGCAGAGCTTCTCGACCGAGGTGGAGCA 864  
Qy 322 eIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAl 342  
Db 865 GCTGTCCAGGTTTCGTACCCAAACATTTGTGACTTCTGCTGCTACTGTGCTCAGAACGG 924  
Qy 342 aLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGlu-----LeuHisLe 360  
Db 925 CTCTTACTGCTGGTGTACGGCTTCTGCCCAACGGCTCCCTGGAGGACCGTCTCCACTG 984  
Qy 360 uHisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySe 380  
Db 985 CCAGACCCAGCCCTGCCCACTCTCTCTGCTGCTCAGCGACTGGACATCTCTTGGGGTAC 1044  
Qy 380 rAlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleHisArgAsp11 400  
Db 1045 AGCCCGGCAATTGATTTTACATCAGGAC---AGCCCCAGCTCATCTCATGAGACAT 1101  
Qy 400 eLysAlaSerAsnIleLeuLeuAspPheLysPheGluAlaLysValAlaAspPheGlyLe 420  
Db 1102 CAAGAGTTTCCAAAGTCTCTTGAGAGAGGCTGACACCCCAAGCTGGGAGACTTTGGCT 1161  
Qy 420 uAlaLysIle-----AlaSerAspThrAsnThrHisValSerThrArg-- 434  
Db 1162 GCGCGGTTTCAGCCGCTTTCGCGGTCCAGCCCGCCAGCAGCAGCATGGTGGCCCGGAC 1221  
Qy 435 -----ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyLysLe 452  
Db 1222 ACAGACAGTGGGGCACCTGGCTTACTTCCCGAGGAGTACATCAAGACGGGAGGCT 1281  
Qy 452 uThrGlyLysSerAspValPheSerPheGlyValValLeuLeuGluLeuIleThrGlyAr 472  
Db 1282 GGCTGTGACACGGACACTTTCAGCTTTGGGTGGTGTAGTGTAGAGACTTGGCTGGTCA 1341  
Qy 472 gArgProVal-----AspAlaAsnValTyrValAspAspSerLeuValAspTr 489  
Db 1342 GAGGCTGTGAAGACGACCGTGGCCAGGACCAAGTATCTGAAAGAC---CTGGTGGAA-- 1396  
Qy 489 pAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAl 509  
Db 1397 -----GAGGAGGCTGAGAGGCTGGAGTGGCTTTGAGAGACCCACGAG 1440  
Qy 509 aLysMetAsnGlyTyrAspArgGluGluMetAlaArgMetValAla----- 525  
Db 1441 CACACTGCAAGCAGTCTGGCTGCAGATGCTGGCTGTCTCCATCGCCATGCAGATCTA 1500  
Qy 525 ----- 525  
Db 1501 CAAGAAGCACCTGGACCCCGGCGCCCTGCCACCTGAGCTGGGCCTGGG 1560  
Qy 526 ----CysAlaAlaAlaCysValArgHisSerAlaArgArgArgProArgMetSerGln1 544  
Db 1561 CCAGCTGGCTGCTGCTGCACCGCGGGGCCAAAGAGGCGCTCTCTATGACCACGGT 1620

Qy 544 eValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArgProGl 564  
Db 1621 GTACAGAGGCTAGAG-----AAGCTGCAGGCAGTGTGGCGGGGTG---CCCGG 1668  
Qy 564 Y-----GlnSerAsnValTyrSerSer-- 571  
Db 1669 GCATTTGGAGGCCCGCCAGCTGCATCCCTTCCCGCAGGAGAACTCCTTACGTGTCCAG 1728  
Qy 572 -----TyrGlyGlySerThrAspTyrAsp----- 579  
Db 1729 CACTGGCAGAGCCACACAGTGGGCTGTCTCATGTCAGCCCTGTCAGCGCCATCAGGAGC 1788  
Qy 580 -SerSerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGl 599  
Db 1789 CAGTGGCCAGCAGCAGCAGCAGCTGCAGAGAGGCCCCCAACAGCCCGGTGGAGAGTGACGA 1848  
Qy 599 uTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSer-- 618  
Db 1849 G-----AGCTAGCGCGGCTCTCTGCTGCGCTGCGCTCCTGGCCTTGACTTCCAGAGTG 1902  
Qy 619 -----GlySerSerSerGluGlyGlnThrThrArgGl 629  
Db 1903 CCCTCTGGACCCAGCACCCCTCAGGAGGCGCGGTCTCTCAGGGGACACGGC-AGGAG 1961  
Qy 629 uMetGluMetGly-----LysIleLysArgThrGlyGlnGlyTyrSerGlyPro 645  
Db 1962 AATCAGCTGGGGAGTGGCCAGGATCCCGGCCACAGCGGTGGAAGGACTGGCCCT 2019

## RESULT 14

US-09-602-472A-5  
; Sequence 5, Application US/09602472A  
; Patent No. 6608240  
; GENERAL INFORMATION:  
; APPLICANT: Bidney, Dennis L.  
; APPLICANT: Hu, Xu  
; APPLICANT: Li, Guihua  
; TITLE OF INVENTION: Sunflower Disease Resistance Genes  
; FILE REFERENCE: 35718/200630  
; CURRENT APPLICATION NUMBER: US/09/602,472A  
; PRIOR FILING DATE: 2000-06-23  
; PRIORITY APPLICATION NUMBER: 60/140,876  
; PRIOR FILING DATE: 1999-06-24  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2114  
; TYPE: DNA  
; ORGANISM: Helianthus annuus  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(2114)  
; OTHER INFORMATION: n = A,T,C or G  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1475  
; OTHER INFORMATION: n = A,T,C or G  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1475  
; OTHER INFORMATION: n = A,T,C or G  
; US-09-602-472A-5

Alignment Scores:  
Pred. No.: 2,96e-16 Length: 2114  
Score: 578.00 Matches: 143  
Percent Similarity: 55.07% Conservative: 85  
Best Local Similarity: 34.54% Mismatches: 136  
Query Match: 16.74% Indels: 51  
DB: 4 Gaps: 15

US-10-086-464-2 (1-647) x US-09-602-472A-5 (1-2114)

```

Db      1056 GAACGATACCTCACAATGTGTGGAGTGAAGACAAATAAGTACACAAGTTTGATTTTATA 1115
Qy      594 AlaLeu-----GlyThrGlnGluTyrAsnAlaThrGlyGluTyrSerAsnProThrSer 611
Db      1116 ACGATTCTTGATCGCAACTGAAAACTTCTCTGAAGCTAATAAGATTTCATCTCTCATAT 1175
Qy      612 AspTyrGlyLeuTyrProSerGlySerSerSerGluGlyGln 625
Db      1176 GAT---TCCGTGTAC---AAGGACGGCTACAAAATGGCAA 1211

RESULT 15
US-08-265-628-1
; Sequence 1, Application US/08265628
; Patent No. 5821094
; GENERAL INFORMATION:
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Goring, Daphne
; TITLE OF INVENTION: S-LOCUS RECEPTOR KINASE GENE IN A
; TITLE OF INVENTION: SELF-INCOMPATIBLE BRASSICA NAPUS LINE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 W. Madison St. Suite 3400
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/265,628
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/959,945
; FILING DATE:
; APPLICATION NUMBER: US 07/847,564
; FILING DATE: 03-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pochopien Ph.D., Donald J.
; REGISTRATION NUMBER: 32,167
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-707-8889
; TELEFAX: 312-707-9155
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2749 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Brassica napus
; STRAIN: oleifera
; INDIVIDUAL ISOLATE: W1
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: S-locus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2574
; PUBLICATION INFORMATION:
; AUTHORS: GORING, DAPHNE
; AUTHORS: ROTHSTEIN, STEVEN J.
; TITLE: THE S-LOCUS RECEPTOR KINASE GENE IN A
; TITLE: SELF-INCOMPATIBLE BRASSICA NAPUS LINE ENCODES A
; TITLE: FUNCTIONAL SERINE/THREONINE KINASE
; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2749

```

## US-08-265-628-1

## Alignment Scores:

Pred. No.: 5,04e-16 Length: 2749  
Score: 575.00 Matches: 154  
Percent Similarity: 49.80% Conservative: 94  
Best Local Similarity: 30.92% Mismatches: 160  
Query Match: 16.65% Indels: 91  
DB: 1 Gaps: 14

US-10-086-464-2 (1-647) x US-08-265-628-1 (1-2749)

```
Qy 130 SerSerAspGlyLeuSerThrGlyValValGlyLeuAlaIleGlyGlyValAlaLeu 149
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1327 AACCGAATGGAAACCATGACGCTTGATTTGGAGTTTGTGTT----- 1371
Qy 150 LeuValIleValThrLeuIleCysLeuLeuCysValLysLysArgArgAspGluGlu 169
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1372 CTGCTTCTTATGATCATGTTCTGC---CTCTGGAAAGAGAAACAAAGCGAGCAAAACA 1428
Qy 170 AspAlaTyrTyrValProProProProGlyProLysAlaGlyGlyProTyrGly 189
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1429 ACTGCAACATCTATTGTAAT----- 1449
Qy 190 GlyGlnGlnGlnGlnTrpArgGlnGlnAsnAlaThrProProSerAspHisValValThr 209
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1450 -----CGACAGAGAAAC-----CAAGATTTCGTAATGAAC 1479
Qy 210 SerLeuProProProProLysAlaProSerProProArgGlnProProProProPro 229
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1480 GGGATG-----ATACTATCAAGCAAGACAGACAGTGCCT----- 1512
Qy 230 ProProPheMetSerSerSerGlySerAspTyrSerAspArgProValLeuProPro 249
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1513 -----ATAGACAACAAACCTGAGGAATTTGGAACCTTCCATTGATA----- 1551
Qy 250 ProSerProGlyLeuValLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAla 269
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1552 -----GAGTTGGAAAGCTGTTGTC 1569
Qy 270 ArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrVal 289
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1570 AAAGCCACCGAAATTTCTCAATTTGACAACTCGGACAAAGTGGTTTCGGTATTGTT 1629
Qy 290 HisLysGlyValLeuProSerGlyLysGluValAlaValLysGlnLeuLysValGlySer 309
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1630 TACAAGGCTAGATTACTTGTATGGGCAAGAAATTCGGTAAAAAGGCTATCAAAAACGTCG 1689
Qy 310 GlyGlnGlyGluArgGluPheGlnAlaGluValGluIleLeuSerArgValHisArg 329
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1690 GTTCAAGGGACTGGTGAGTTTATGAATGAGGTGAGATTGATCGCGAGGCTTTCAGCATATA 1749
Qy 330 HisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGlu 349
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1750 AACCTGTCCGAATTTCTGGCTGTTGATTGAGCAGACAGAGAAGATGCTGTTATATGAG 1809
Qy 350 PheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGlyGluGlyArg---ProThrMet 368
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1810 TATTTAGAAAATTTAAGGCTCGATTCTTATCTCTTCGGAATAAAACGAAGCTCTACGTTA 1869
Qy 369 GluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLysGlyLeuSerTyrLeuHis 388
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1870 AATTGGGAAGGACAGATTCAACATTACCAATGGTGTGCTCGAGGACTTTTATATCTTCAT 1929
Qy 389 GluAspCysAsnProLysIleIleHisArgAspIleLysAlaSerAsnIleLeuIleAsp 408
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1930 CAAGACTCACGGTTAGGATTAATCAACAGATATGAAGTAAGTAACATTTTCCTTGAT 1989
Qy 409 PheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLysIleAlaSerAspThrAsn 428
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 1990 AAAAATATGACACCAAGATCTCGGATTTTGGGATGGCCAGAACTTTTGCAAGGGACGAG 2049
Qy 429 ThrHisValSerThrArg---ValMetGlyThrPheGlyTyrLeuAlaProGluTyrAla 447
```

```
Db 2050 ACTGAAGCTAACACAGGAAGGTGTCGGAACCTTACGGCTACATGTCCTCCGGATACGCA 2109
Qy 448 AlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGlyValValLeuLeuGlu 467
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 2110 ATGGATGGGGTATTCTCGGAAAAAATCAGATGTTTTCAGTTCATGTTTGGAGTCATTGTTCTGAA 2169
Qy 468 LeuIleThrGlyArgArgProValAspAlaAsnValTyrValAspAspSerLeuVal 487
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 2170 ATTGTTAGTGAAGAAAGGAACAGAGGATTCTACAACCTTGAACACCAAGAAAAAATCTCTTA 2229
Qy 488 Asp-----TrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp----- 502
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 2230 AGCTATGATGAGTCTACGCGAGGGAAGCGCTGAGAAATTTGTATCCAGTCAATC 2289
Qy 503 PheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMetAlaArg 522
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 2290 GTAGATTTCATTGTCATCATTTACCAGCAACC-----TTTCAACCAAGAAAGTTCTAAAA 2343
Qy 523 MetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgArgProArgMetSer 542
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 2344 TGCATACAAAATTGGTCTCTTGTGTGTTCAAGAACGTCGAGAGCATAGACCAACCATGTGCG 2403
Qy 543 GlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArg 562
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 2404 TCCGTGTTTGGATGCTTGGAAAGTGAAGCA-----ACAGAGATTCTGTGAGCCTACACCG 2457
Qy 563 ProGlyGlnSerAsnValTyrSerSerTyrGlyGlySer-----ThrAspTyrAspSerSerGln 582
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 2458 CCAGTTTATTCCTCGGAAGAGTCTCTTATGAAAATAATCTTTCATCAAGTAGACATTGC 2517
Qy 583 TyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
    :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
Db 2578 TACGA-AATCCGTTGAGAAAGTTTCAGATAAATTAACATAATGGGGTGACCGGATAT 2630
```

Search completed: May 13, 2004, 01:36:26

Job time : 187 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 13, 2004, 00:33:06 ; Search time 677 Seconds  
(without alignments)  
4329.253 Million cell updates/sec

Title: US-10-086-464-2  
Perfect score: 3453  
Sequence: 1 MSSAPSGTGSPPSPNSNT.....REMEGKIKRTGGYSGPSL 647

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPRO.spool.p/US10086464/runat\_11052004\_121817\_7949/app\_query.fasta\_1.839  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=45 -DALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10086464 @CGN 1 1 333 @runat\_11052004\_121817\_7949  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Result No. Score Match Length DB ID Description

†  
SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

1	3453	100.0	1944	14	US-10-086-464-1	Sequence 1, Appli
2	3453	100.0	2189	14	US-10-086-464-3	Sequence 3, Appli
3	2985.5	66.2	2451	13	US-10-424-599-95479	Sequence 95479, A
4	1870	54.2	1724	13	US-10-425-114-2326	Sequence 2326, Ap
5	1833	53.1	2880	13	US-10-425-114-33030	Sequence 33030, A
6	1833	53.1	2881	13	US-10-425-114-33031	Sequence 33031, A
7	1748.5	50.6	1902	14	US-10-086-464-10	Sequence 10, Appl
8	1748.5	50.6	1939	14	US-10-086-464-9	Sequence 9, Appli
9	1737.5	50.3	1424	13	US-10-425-114-9117	Sequence 9117, Ap
10	1687	48.9	2025	14	US-10-086-464-13	Sequence 13, Appl
11	1687	48.9	2104	14	US-10-086-464-12	Sequence 12, Appl
12	1539	44.6	2655	13	US-10-424-599-57846	Sequence 57846, A
13	1539	44.6	2655	13	US-10-425-114-7339	Sequence 7339, Ap
14	1437.5	41.6	2190	13	US-10-425-114-14127	Sequence 14127, A
15	1437	41.6	1500	13	US-10-425-114-6300	Sequence 6300, Ap
16	1437	41.6	1552	13	US-10-424-599-47702	Sequence 47702, A
17	1415.5	41.0	2196	14	US-10-086-464-7	Sequence 7, Appli
18	1415.5	41.0	2231	14	US-10-086-464-6	Sequence 6, Appli
19	1400.5	40.6	2261	14	US-10-086-464-15	Sequence 15, Appl
20	1399	40.5	2196	14	US-10-086-464-16	Sequence 16, Appl
21	1379	39.9	2508	13	US-10-425-114-25269	Sequence 25269, A
22	1335.5	38.7	1585	13	US-10-424-599-71008	Sequence 71008, A
23	1239	35.9	1588	13	US-10-425-114-35509	Sequence 35509, A
24	1238.5	35.9	1462	13	US-10-425-114-29328	Sequence 29328, A
25	1233.5	35.7	1134	13	US-10-425-114-36420	Sequence 36420, A
26	1210	35.0	1795	13	US-10-424-599-42232	Sequence 42232, A
27	1116.5	32.3	1179	13	US-10-425-114-4027	Sequence 4027, Ap
28	1116.5	32.3	1186	13	US-10-425-114-4273	Sequence 4273, Ap
29	1115.5	32.3	981	13	US-10-425-114-7549	Sequence 7549, Ap
30	921	26.7	729	13	US-10-425-114-12191	Sequence 12191, A
31	892.5	25.8	760	13	US-10-425-114-12314	Sequence 12314, A
32	866.5	25.1	951	13	US-10-425-114-1922	Sequence 1922, Ap
33	860.5	24.9	1777	13	US-10-425-114-24475	Sequence 24475, A
34	859.5	24.9	1817	13	US-10-425-114-30665	Sequence 30665, A
35	840	24.3	2025	13	US-10-425-114-3666	Sequence 3666, Ap
36	838.5	24.3	3148	16	US-10-260-238-465	Sequence 465, App
37	837	24.2	2284	13	US-10-425-114-3455	Sequence 3455, Ap
38	825	23.9	2012	13	US-10-424-599-29210	Sequence 29210, A
39	819.5	23.7	1011	13	US-10-425-114-13131	Sequence 13131, A
40	816.5	23.6	2124	9	US-09-938-842A-1079	Sequence 1079, Ap
41	816.5	23.6	2124	11	US-09-938-842A-1079	Sequence 311, App
42	811	23.5	3453	15	US-10-259-165-311	Sequence 25617, A
43	806.5	23.4	2063	13	US-10-424-599-25617	Sequence 33063, A
44	804	23.3	2374	13	US-10-425-114-33063	Sequence 5, Appli
45	802.5	23.2	2011	16	US-10-168-844-5	

ALIGNMENTS

RESULT 1  
US-10-086-464-1  
; Sequence 1, Application US/10086464  
; Publication No. US20020199218A1  
; GENERAL INFORMATION:  
; APPLICANT: GORING, Daphne R. et al.  
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES  
; FILE REFERENCE: P 25,762-A USA  
; CURRENT APPLICATION NUMBER: US/10/086,464  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 10/069,304  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: PCT/CA00/009566  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/149,466  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: US 60/159,122  
; PRIOR FILING DATE: 1999-10-13  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1944  
; TYPE: DNA

```
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1944)
US-10-086-464-1

Alignment Scores:
Pred. No.:          9,36e-216      Length:      1944
Score:             3453.00         Matches:      647
Percent Similarity: 100.00%        Conservative: 0
Best Local Similarity: 100.00%      Mismatches:   0
Query Match:       100.00%         Indels:       0
DB:                14              Gaps:        0

US-10-086-464-2 (1-647) x US-10-086-464-1 (1-1944)

Qy   1 MetSerSerAlaProSerProGlyThrGlySerProProSerProProSerAsnSerThr 20
Db   1 ATGTCCTCGGCGCGCTCTCCGGGACTGGTTGGCTCCCATCTCCACCATCAAACTCCACA 60

Qy   21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40
Db   61 ACCACCCTCTCTCCAGCTTCGGCTCTCTCTCCACCCACACCTCTCTCTCTCTCCGCG 120

Qy   41 ProSerThrIleProThrSerProProProProSerSerArgSerThrProSerAlaProPro 60
Db   121 CCATCCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

Qy   61 ProSerProProThrProSerThrProGlySerProProProProLeuProGlnProSerPro 80
Db   181 CCATCTCCACCACTCCATCTACGCGGGAGTCTCCACCTCTCTCTCTCTCTCTCTCTCTCA 240

Qy   81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100
Db   241 CCGCTCCAACTACGCGCGGATCTCCACCCGACCTGTTACTCTCTCTCTCTCTCTCTCTCT 300

Qy   101 ProProSerValProGlyProProSerAsnProSerArgGluGlyGlySerProArgPro 120
Db   301 CCACCTTCAGTCCGAGGACCACTGCTCCAAATCTCTCAACGGAAGGAGGATCTCTCTCGACCT 360

Qy   121 ProSerSerProSerProProSerProSerSerAspGlyLeuSerThrGlyValValVal 140
Db   361 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420

Qy   141 GlyIleAlaIleGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160
Db   421 GGAATCGCCATCGGAGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480

Qy   161 LysLysLysArgArgArgAspGluAspAlaTyrTyrValProProProProProProPro 180
Db   481 AAGAAGAAACACGCGAGAGACGAAGAAGATCTTACTATATGTTCTCTCGCCCACTCTCTCT 540

Qy   181 GlyProLysAlaGlyProTyrGlyGlyGlnGlnGlnGlnTyrArgGlnGlnAsnAla 200
Db   541 GGTCCCAAGCCGAGGACCTTACGGTGGACAGCAGCAACAATGGCGCAACAACAAACGCA 600

Qy   201 ThrProProSerAspHisValValThrSerLeuProProProProLysAlaProSerPro 220
Db   601 ACACCACGTCAGATCATGTCTGAGGTCACCTACCACCACTTAAGGCTCCATCTCCA 660

Qy   221 ProArgGlnProProProProProProProProProPheMetSerSerSerGlyGlySerAsp 240
Db   661 CCACGGCAACCTCTCTCACTCCACCCGCTTTCATGACGACGCGCGGCTCCGAC 720

Qy   241 TyrSerAspArgProValLeuProProProProProGlyLeuValLeuGlyPheSerLys 260
Db   721 TACTCGGACCGTCCAGTCTCTCTCTCCACCTCTCCAGGCTTGTGTAGGCTTCTCCAA 780

Qy   261 SerThrPheThrTyrGluGluLeuAlaAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeu 280
Db   781 AGCATTTTCATACGAGGAGTAGTAGAGCCACCAATGTTCTCTCGGCGGCAACTTG 840

Qy   281 LeuGlyGlnGlyPheGlyTyrValHisLysGlyValLeuProProSerGlyLysGluVal 300

841 TTAGCAACGCGGTTCCGTTACGTGCACAAAGGTGTGTTCCTAGTGGGAAGAAGTT 900
Qy   301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320
Db   901 GCTGTGAAGCAGTTGAAGTTGGAGTGGTCAGGAGAGAGGAGTTTCAGCAGAGGTT 960

Qy   321 GluIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAla 340
Db   961 GAGATCATCAGCAGATTCCACACAGCATCTGGTGTCTCTTGTGTGGTATTATGATCGCC 1020

Qy   341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuLeuHisLeu 360
Db   1021 GGTGCCAAAGATTGCTTGTCTATGAGTTGTTCTTAACAACAATCTCGAGCTTCACCTC 1080

Qy   361 HisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySer 380
Db   1081 CATGCCGAGGACGCGCTTACATGGAATGGAGCACACAGATTGAAGATTGCTCTTGGATCT 1140

Qy   381 AlalysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400
Db   1141 GCTAAAGACTTTCTTATCTTCATGAAGATTGCAATCTCTAAATCAATTCACCGTGATATC 1200

Qy   401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420
Db   1201 AAGCTTCAAAACATATTGATAGATTTCAGTTTGAAGCTAAGGTTCGTGATTTTGGTCTT 1260

Qy   421 AlalysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440
Db   1261 GCTAAGATTGCTTCTGATACAAACACGATCTATCAACGCTGTGATGGGAACCTTTGGG 1320

Qy   441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460
Db   1321 TACTTGGCTCCGAAATACGCTCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTCA 1380

Qy   461 PheGlyValValLeuLeuLeuLeuLeuThrGlyArgArgProValAspAlaAsnVal 480
Db   1381 TTTGGCGTGTGCTTTTGGAGCTCATTTACTGGAGCTGAGCCGCTGATGCGCAACAATGTC 1440

Qy   481 TyrValAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGln 500
Db   1441 TATGTAGATGACAGCTTAGTTGACTGGGCACGACCAATGCTTTAACCGGACATCTGAGCA 1500

Qy   501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
Db   1501 GGAGACTTTGAGGGTTTAGCTGATGCAAGATGAATAATGGGTATGACAGAGAGAGATG 1560

Qy   521 AlaArgMetValAlaCysAlaAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db   1561 GCTCGCATGGTTGCTTGTGCTGGGCTGTGTTGCGCATTCAGCTCGCCGACACCTCGC 1620

Qy   541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db   1621 ATGAGCCAGATTGCGTGGCTTAGAAGGAAATGTATCACTGTCTGATCTTTAACGAGGG 1680

Qy   561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db   1681 ATGAGACAGGTCAAAGCAATGTATACAGCTCATACGAGGAGAACCCGATTATGACTCG 1740

Qy   581 SerGlnTyrAsnGluAspMetLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db   1741 AGCAGTACATTAAGACATGAAGAGTTTAGGAAATTTGGCACTTTGGAACTCAAGAGTAC 1800

Qy   601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db   1801 AACGCCACGCGGTAGTACAGTAAATCCGACCATGACTATGAGCTGTACCCGCTCTGGTTCA 1860

Qy   621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640
Db   1861 AGCAGCGAGGCGCAACACACACGCAATGGAGATGGGGAAGATTAAGAGAACCCTGTCAG 1920

Qy   641 GlyTyrSerGlyProSerLeu 647
```

Db 1921 GGTATAGTGGACCTTCTT 1941

RESULT 2

US-10-086-464-3

; Sequence 3, Application US/10086464

; Publication No. US20020199218A1

; GENERAL INFORMATION:

; APPLICANT: GORING, Daphne R. et al.

; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES

; FILE REFERENCE: P 25,762-A USA

; CURRENT APPLICATION NUMBER: US/10/086,464

; CURRENT FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 10/069,304

; PRIOR FILING DATE: 2002-02-19

; PRIOR APPLICATION NUMBER: PCT/CA00/00966

; PRIOR FILING DATE: 2000-08-18

; PRIOR APPLICATION NUMBER: US 60/149,466

; PRIOR FILING DATE: 1999-08-19

; PRIOR APPLICATION NUMBER: US 60/159,122

; PRIOR FILING DATE: 1999-10-13

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 2189

; TYPE: DNA

; ORGANISM: Brassica napus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2189)

; OTHER INFORMATION:

US-10-086-464-3

Alignment Scores:

Pred. No.: 1.05e-215 Length: 2189

Score: 3453.00 Matches: 647

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 14 Gaps: 0

US-10-086-464-2 (1-647) x US-10-086-464-3 (1-2189)

Qy 1 MetSerSerAlaProSerProGlyThrGlySerProProSerProSerAsnSerThr 20

Db 97 ATGTCTCGCGCGCTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 156

Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerProProPro 40

Db 157 ACCACCACTCTCTCCAGCTTCGGCTCCTCTCCACCAACACACCTCTCTCTCCGCGC 216

Qy 41 ProSerThrThrProThrSerProProProSerSerSerSerThrProSerAlaProPro 60

Db 217 CCATCCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276

Qy 61 ProSerProProThrProSerThrProGlySerProProProProProGlnProSerPro 80

Db 277 CCATCTCCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 336

Qy 81 ProAlaProThrThrProGlySerProProAlaProValThrProProThrArgAsnPro 100

Db 337 CCGCTCCCACTACGCGCGGATCTCCACCGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 396

Qy 101 ProProSerValProGlyProProProSerAsnProSerArgGluGlyGlySerProArgPro 120

Db 397 CCACCTTCAGTCCCGGACCAACCGCTCCCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 456

Qy 121 ProSerSerProSerProProSerProSerSerSerSerSerSerThrGlyValValVal 140

Db 457 CCATCT 516

Qy 141 GlyIleAlaIleGlyGlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCys 160

Db 517 GGAATCGCCATCGGAGGAGTCGCT 576

Qy 161 LysLysLysArgArgArgAspGluGluAspAlaTyrTyrValProProProProProPro 180

Db 577 AAGAGNAAACGACGGAGACGAGAGAGATGCTTACTATGTTCTCTCGCCACCTCTCTCTCT 636

Qy 181 GlyProLysAlaGlyGlyProTyrGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200

Db 637 GGTCCCAAGCCGAGGACCTTACGGTGGACAGCAGCAACAATGGCGGCAACAAGCGCA 696

Qy 201 ThrProProSerAspHisValValThrSerLeuProProProProProProProProPro 220

Db 697 ACACCAACCGTCAGATCATGTCGTGACGTCACTACCAACCAACCACTAAGGCTTCCATCTCCA 756

Qy 221 ProArgGlnProProProProProProProPheMetSerSerSerSerGlyGlySerAsp 240

Db 757 CCACGGCAACTCTCTCCACCTCCACCAACCGCTTCTCATGACGACGACGCGGCTCCGAC 816

Qy 241 TyrSerAspArgProValLeuProProProProProProGlyLeuValLeuGlyPheSerLys 260

Db 817 TACTCGGACCGTCCAGTCT 876

Qy 261 SerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGlyAlaAsnLeu 280

Db 877 AGCACTTTTCATACGAGGAGCTAGCTAGACCAACCAATGGTCTCTCTCGAGGCGCACTTG 936

Qy 281 LeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluVal 300

Db 937 TTAGGACNAGCGGGTTCGGTTACGTGCACAAAGTGTGTCTCTGTGGTATTTGATGCAAGAGTT 996

Qy 301 AlaValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluVal 320

Db 997 GCTGTGAAGCAGTTGAAAGTTGGAGTGGTTCAGGAGAGGAGGAGTTTCAGGAGAGAGTT 1056

Qy 321 GluIleSerArgValHisArgHisLysValSerLeuValSerLeuValGlyTyrCysIleAla 340

Db 1057 GAGATCATACGAGAGTTTCCACCAACGAGTCTGTGTCTCTCTGTGGTATTTGATGCAAGAGTT 1116

Qy 341 GlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnLeuLeuGluLeuHisLeu 360

Db 1117 GGTGCCAAAGATGCTTGTCTATGAGTTTGTCTTAAACAACATCTCTCGAGCTTCACTCTC 1176

Qy 361 HisGlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysIleAlaLeuGlySer 380

Db 1177 CATGGCGAGGACGGCTCAATGGAATGGAGACCAACAGATTTGAAGATTTGCTCTCTGGATCT 1236

Qy 381 AlaLysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIle 400

Db 1237 GCTAAAGGAGCTTTCTTATCTTCAAGATTTGCAATCTTAAATCATTTCCCGGTGATATC 1296

Qy 401 LysAlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeu 420

Db 1297 AAGCTTCAAAACATATTTAGATAGATTTTCAAGTTTGAAGCTTGAAGTTGCTGATTTTGGTCTT 1356

Qy 421 AlaLysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGly 440

Db 1357 GCTAAGATTTCTCTGATCAAAACACGATGATCAACAGTGTGATGGGAACCTTTTGGG 1416

Qy 441 TyrLeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSer 460

Db 1417 TACTTGGCTCTCGGAATACGCTGCAAGCGGAAGCTCACGGAGAAGTCTGACGTTTCTCTCA 1476

Qy 461 PheGlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnAsnVal 480

Db 1477 TTTGGGGCTGTGCTTTTGGAGCTCATTTACTTGGAGCTGAGCTGAGCTTGTATGCAACAAATGTC 1536

Qy 481 TyrValAspAspSerLeuValAspTyrAlaArgProLeuLeuAsnArgAlaSerGluGln 500

Db 1537 TATGTAGATCACAGCTTAGTTGACTGGGCACGACCATTTGCTTAACCGGACATCTGACCAA 1596

Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520

Db 1597 GGAGACTTTGAGGCTTGTGATGCAAAAGATGAATAATGGTATGATGACAGAGAGATG 1656

```
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArg 540
Db 1657 GCTCGCATGGTGTGTGCTGCGGCTGTGTTTCGCCAATTCAGCTCGCGCAGACCTCGC 1716
Qy 541 MetSerGlnIleValArgAlaLeuGluCysValSerLeuSerAspLeuAsnGluGly 560
Db 1717 ATGAGCCAGATTGCGTGTAGAGGAATGTATCACTGTCACTTAAAGAGGG 1776
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSer 580
Db 1777 ATGAGCAGGTCAAAGCAATGTATACAGCTCATACGAGGAGGACCGATTATGACTCG 1836
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1837 AGCCAGTACAAATGAAGACATGAAGAAGTTTAGGAAAATGGCACATGGAACATCAAGAGTAC 1896
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1897 AACGCCACGGGTGAGTACAGTAATCCGACCACTGACTATGGACTGTACCCGCTGTGTTCA 1956
Qy 621 SerSerGluGlyGlnThrThrArgGluMetGluMetGlyLysLysLysArgThrGlyGln 640
Db 1957 AGCAGCGAGGCGCAACACACACGGAATGGAGATGGGAAGATTAAAGAGAACCGGTGAG 2016
Qy 641 GlyTyrSerGlyProSerLeu 647
Db 2017 GGTATATAGTACCTTCTCT 2037

RESULT 3
US-10-424-599-95479
; Sequence 95479, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95479
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57229C.1
US-10-424-599-95479

Alignment Scores:
Pred. No.: 9,23e-140 Length: 2451
Score: 2285.50 Matches: 450
Percent Similarity: 80.86% Conservative: 74
Best Local Similarity: 69.44% Mismatches: 97
Query Match: 66.19% Indels: 27
DB: 13 Gaps: 16

US-10-086-464-2 (1-647) x US-10-424-599-95479 (1-2451)

Qy 13 ProSerProSerAsnSerThrThrThrThrProProProAla-----SerAlaPro 30
Db 246 CCCCTCCCTCGGCTC---GCAACTCCGTCGTCAGCTCCGCGGCAACTCCGCTCGTCTCCT 302
Qy 31 ProProThrThrProSerProProProSerThrThrThrProProProPro 50
Db 303 CCGCGCGGACACCTTCTTCGCCACCTCCG-----TCAACTCTCTTCTCCGCCCTCCG 356
Qy 51 SerSerArgSerThrProSerAlaProProProProProProProProGly 70
Db 357 -----GGACTCCCTCCGCTCTCCACCGTCCACTCTCTCTGCTTCTCTCTCCACCG 407
```

```
Qy 71 SerProProProLeuProGlnProSerProProAlaProThrThrProGlySerProPro 90
Db 408 TCCACTCAACACCGCGCTCAACTTTCGCGCCATCG---ACTTCTCCGCGTCCCGGCCA 464
Qy 91 -----AlaProValThrProThrArgAsnProProProProSerValProGlyProPro 108
Db 465 TCGCACTCGCGCGCTCGCTCCAGTGGCGGCGGAGCGCAGCACTCCGAGTCCACCG 524
Qy 109 SerAsnProSerArgGluGlyGlySerProArgProProSerSerProSerProSer 128
Db 525 TCCGGAGCTCGGCTCTCCTCGCTCGGATTCGAGACCGACCACTCCGCTCGCTCCGTCG 584
Qy 129 ProSerSerAspGlyLeuSerThrGlyValValGlyLysLysLysLysLysLysLys 148
Db 585 TCTTCTCTCGTGGATATTCGACCGGTGTGGTGGGAATCGCCGTTGGGGCTGTGGCG 644
Qy 149 LeuLeuValIleValThrLeuIleCysLeuLeuCys---LysLysLysLysLysLysLys 167
Db 645 GTTCTTCTGTGTGAGCAATCTCTGCATATGTTGCCGGAAGAAGAAGACGCTGAT 704
Qy 168 GluGluAspAlaTyrTyrValProProProProProPro---GlyProLysAlaGlyGly 186
Db 705 GAAGAG-----TACTATGCTCGCGCGCGCAACCGCGCGGCGGACCTAAA---GATGAT 755
Qy 187 ProTyrGlyGlnGlnGlnGlnTyrArgGlnGlnAsnAlaThrProProSerAspHis 206
Db 756 GCATATGGTGGTCCCGCACGTCAATGG---CAACACAAATGTTCCCTCTCTCAAGATCAT 812
Qy 207 ValValThrSerLeuProPro-----ProProLysAlaProSerProProArgGln 223
Db 813 GTGGTCTCAATGATGCTCTCAAGCCATCGCCACCACTGCTCCACCGGCTTATGCTGGT 872
Qy 224 ProProProProProProProProPheMetSerSerSerGly-----GlySerAspTyr 241
Db 873 CAACCTCCCGCACCGCTCTCTTTCATCAGCAGCAGTGGCGGATCTGGATCAAACTAT 932
Qy 242 SerAspArgProValLeuProProProProGlyLeuValLeuGlyPheSerLysSer 261
Db 933 TCAGCGCGTGAAATTTCTCTCTCTCTCTCCAGAAATTCATTTGGGATTTCTTAAGAGC 992
Qy 262 ThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeu 281
Db 993 ACATTACGATGAGGAGTTGGCAGCGCAACTGATGCTCTCTGATGCCAACCTCTCT 1052
Qy 282 GlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAla 301
Db 1053 GGACAGGAGGATTTGGATATGTGCACAGAGGAATTTCTCCCAACGCGCAGGAGGTGCA 1112
Qy 302 ValLysGlnLeuLysValGlySerGlyGlnGluArgGluPheGlnAlaGluValGlu 321
Db 1113 GTGAAGCAATTTGAAGGCTGGAAGCGGCGCAGGGAGCGTGAATTCCAAGCTGAAGTTGAG 1172
Qy 322 IleIleSerArgValHisHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGly 341
Db 1173 ATAATTAGCGGTGCTCCATCACAAGCATCTTGTTCTTTGGTTGGTACTGCTCACTGGG 1232
Qy 342 AlaLysArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHis 361
Db 1233 TCCGAGAGGCTGCTTGTATTGAATTTGTTTCCCAACACACATTTGGAATTCATTTGTCAT 1292
Qy 362 GlyGluGlyArgProThrMetGluTyrSerThrArgLeuLysLysLysLysLysLysLys 381
Db 1293 GGAAGAGGCGGACCTTACATGATTTGGCCCAAGATGATCATCTTAAGATCATCTCGTATCAA 1352
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysLysLysLysLysLysLys 401
Db 1353 AAGGACTGCGGTATCTTCATGAGATTGTCATCTTAAGATCATCTCGTATCAA 1412
Qy 402 AlaSerAsnLeuLeuLeuAspPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 1413 GCTGCCAACATCTCTCTGATTTTAAAGTTTGAAGCAAGGTTGAGATTTTGGTCTTGCA 1472
Qy 422 LysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyr 441
```





```
Qy 483 AspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502
Db 2152 GATGACAGCTTGGTTGACTGGCGAAGCCATTACTGATCGAGCATTGAGGATGGTAA 2211
Qy 503 PheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMetAlaArg 522
Db 2212 TATGATGCTTTAGTGGTACTCTCGCTGGGAAGACTTCAATCCTAATGATGATGGCAAGA 2271
Qy 523 MetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArgMetSer 542
Db 2272 ATGATACGCTGTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2331
Qy 543 GlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGlyMetArg 562
Db 2332 CAGGTCTGCTGGGCTTTGGAGGCAATGTCTTTGGAGGACCTTAATGAAGGTGTTGCG 2391
Qy 563 ProGlyGlnSerAsnValTyrSerSerTyrGlyGlySerThrAspTyrAspSerSerGln 582
Db 2392 CTTGGCCATAGCCGCTTCTTTGGGTCTATC---AGCAGCTCCGATTACGATTCTGGCCAG 2448
Qy 583 TyrAsnGluAspMetLysValPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAla 602
Db 2449 TACACGAGGACATGAAGAAGTTCAAGAAGATGGCATTCACAAACAC-----TATACC 2502
Qy 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSer 622
Db 2503 AGCAGCCAAATACAGCGCGCCACACCAAGTGAATATGACAGATACCGTCTGCATCAAGCAGC 2562
Qy 623 GluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly----- 639
Db 2563 GAGGCGCCACAGCAGCAGGAGATGAGTGGGTGCAATGAAGAAAGTGGTCTACAGTGGT 2622
Qy 640 -----GlnGlyTyrSerGlyProSer 646
Db 2623 GGCTACAGCTCAGGATACAGCGGAGCCTCG 2652

RESULT 6
US-10-425-114-33031
; Sequence 33031, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 33031
; LENGTH: 2881
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17055B12_FLI
; US-10-425-114-33031

Alignment Scores:
Pred. No.: 2,81e-110 Length: 2881
Score: 1833.00 Matches: 356
Percent Similarity: 79.80% Conservative: 51
Best Local Similarity: 69.80% Mismatches: 85
Query Match: 53.08% Indels: 18
DB: 13 Gaps: 9

US-10-086-464-2 (1-647) x US-10-425-114-33031 (1-2881)
Qy 150 LeuValIleValThrLeuIleCysLeuLeuCys-----LysLysLysArgArgAsp 167
```

```
Db 1139 GTGTCGTCGCCAGCTTCTATCTGCTGCTGCTCGCAAGAAGCGCGCGCGCGCG 1198
Qy 168 GluLysAspAlaTyrTyrValProProProProProGlyProLysAlaGlyGlyPro 187
Db 1199 CCGCGCGCTCAGCAGCTACGATACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1255
Qy 188 TyrGlyGlyGlnGlnGlnGlnTyrArgGlnGlnAsnAla-----ThrProProSerAsp 205
Db 1256 TATGGTGGACGTACACAGATTGG---CAGCAAAATGCGCTCTGCTGCCACCCCTGAA 1312
Qy 206 HisValValThrSerLeuPro---ProProProLysAlaProSerProProArgGlnPro 224
Db 1313 CATGTGTCGAAGTGCACCTTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1372
Qy 225 ProProProProProProProPheMetSerSerSerGly-----GlySerAspTyrSer 242
Db 1373 CCAGCGCGCGCACACAGCGATGCTAAATAGTAGTGGTGGATCTGGTTCTTAATACTCT 1432
Qy 243 AspArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerLysSerThr 262
Db 1433 GGTGGCGAGATCTACTCCACCATCCCTGGTGTCTTCTTGGCTTCTCGAAGAGCACA 1492
Qy 263 PheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGly 282
Db 1493 TTCACCTTACGAGAGCTGTTGAGGCTTACTGATGATTCGATGCTGATTAATCTCTTGA 1552
Qy 283 GlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAlaVal 302
Db 1553 CAAGTGGTGTGCTGATGTTACAGAGGATGCTGCTTAATGCAAGAGATGCTGTGA 1612
Qy 303 LysGlnLeuLysValGlySerGlyGlnGlyArgGluPheGlnAlaGluValGluIle 322
Db 1613 AACAATTGAATGGGAGTGGCAGGAGAGCGTGAGTCCAGGCTGAGTTGAGATT 1672
Qy 323 IleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGlyAla 342
Db 1673 ATCAGCGGAGTACATCACAAACACCTGTCTTGTGGTGTATGCTTCTTGTGGAGGC 1732
Qy 343 LysArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeuGly 362
Db 1733 AAGAGCTGCTGCTGATGATGTTGTCCCAATTAACACATGGAATTCACATTCATCGC 1792
Qy 363 GluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAlaLys 382
Db 1793 AAGGTGCGACCAACATGGAGTGGCTGTAGATTAAGATCAGTTGGTGGTCCCAAG 1852
Qy 383 GlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAla 402
Db 1853 GGTTAGCTTATCTTCATGAAGACTGCGCATCCAAGATCATCATCGTGCATATAAGGCA 1912
Qy 403 SerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAlaLys 422
Db 1913 TCTAACATCTCTTCTGACTTCCAAATTTGAAGCTTAAGGTTCTGACTTTGGACTTGAAG 1972
Qy 423 IleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyrLeu 442
Db 1973 TTCACTACTGATTAACAACACCCATGTTTCACAAGAGTAAATGGCACCTTTGGGTATTG 2032
Qy 443 AlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPheGly 462
Db 2033 GCACCTGAGTATGATCTCTGCAAGCTAACAGAAAAATCCGATGTATTTCTCTCGGA 2092
Qy 463 ValValIleLeuLeuIleThrGlyArgArgProValAspAlaAsnValTyrVal 482
Db 2093 GTATGCTTCTTCTGAGCTTATTAATGCGCGCGCGCGCGCGCGCGCGCGCGCG 2152
Qy 483 AspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAlaSerGluGlnGlyAsp 502
Db 2153 GATGACAGCTTGGTTGACTGGGCAAGGCCATTACTGATGCGAGCAGCTTGGAGTGGTAA 2212
Qy 503 PheGluGlyLeuAlaAspAlaLysMetAsnGlyTyrAspArgGluGluMetAlaArg 522
```

Db 2213 TATGATGCTTTAGTGATCTCGCTGGGAAGAGACTTCAATCTAATGAGATGGCAAGA 2272  
Qy 523 MetValAlaCysAlaAlaCysValArgHisSerAlaArgArgProArgMetSer 542  
Db 2273 ATGATAGCCTGTGCAGCTGCTGTGTAGCGCCATTCTGCACGCTCGTCCGCCACGATGAGT 2332  
Qy 543 GlnIleValArgAlaLeuGluGluCysValSerLeuSerAspLeuAsnGluGlyMetArg 562  
Db 2333 CAGTCGTTCGGGCTTTGGAGGGAATGTGCTTTGGAGGACCTTAATGAAGGTGTTGG 2392  
Qy 563 ProGlyGlnSerAsnValTyrSerTyrGlyGlySerThrAspTyrAspSerGln 582  
Db 2393 CCTGGCCATAGCCGCTCTTTGGGTGATAC---AGCAGCTCCGATTAGGATTCGGCCAG 2449  
Qy 583 TyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAsnAla 602  
Db 2450 TACAACGAGGACATGAAGAGTTCAGAGAAGATGGCATTCAACAACAAC-----TATACC 2503  
Qy 603 ThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySerSerSer 622  
Db 2504 AGCAGCCAATACAGCGCGCCCAACAGTGAATATGGACAGATACCGTCTGCATCAAGCAGC 2563  
Qy 623 GluGlyGlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGly----- 639  
Db 2564 GAGGCCACACAGACGAGGAGATGGAGTCGGGTGCAATGAAGAAGGTGGCTACAGTGGT 2623  
Qy 640 -----GlnGlyTyrSerGlyProSer 646  
Db 2624 GGCTACAGCTCAGGATACAGCGGAGCCTCG 2653

## RESULT 7

US-10-086-464-10  
; Sequence 10, Application US/10086464  
; Publication No. US20020199218A1  
; GENERAL INFORMATION:  
; APPLICANT: GORING, Daphne R. et al.  
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES  
; FILE REFERENCE: P 25,762-A USA  
; CURRENT APPLICATION NUMBER: US/10/086,464  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 10/069,304  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: PCT/CA00/00966  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/149,466  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: US 60/159,122  
; PRIOR FILING DATE: 1999-10-13  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 1902  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1902)  
US-10-086-464-10

## Alignment Scores:

Pred. No.: 5,88e-105 Length: 1902  
Score: 1748.50 Matches: 367  
Percent Similarity: 69.67% Conservative: 81  
Best Local Similarity: 57.08% Mismatches: 141  
Query Match: 50.64% Indels: 55  
DB: 14 Gaps: 16

US-10-086-464-2 (1-647) x US-10-086-464-10 (1-1902)

Qy 3 SerAlaProSerProGlyThrGlySerProSerProSerAsnSerThrThr 22  
Db 19 TCTGCTCTCCACACAAATCCACCTCTTCTCCATCTCCACCGTCTAATACCAATTCAACC 78

Qy 23 ThrProProAlaSerAlaProProProThrThrProSerSerProProProSer 42  
Db 79 ACCTCTTCTCCG-----CGGCTCCGTCCTCTCTCTCTCTCTCTCTCTCAACGGA 129  
Qy 43 ThrIleProThrProProSerSerArgSerThrProSerAlaProProProSer 62  
Db 130 GACTCATCATCGCCACCTCTCTGATTCACATCTCCACCAAGCTCCACAAGCTCTCTAAC 189  
Qy 63 ProProThrProSerThrProGlySerProProProLeuProGlnProSerProProAla 82  
Db 190 CCTCCTAAT--TCTCTAATAACTCTCTCTCCCTCCGTCACAGGCGGTGGAGAGAA 246  
Qy 83 ProThrThrProGlySerProAlaProValThrProProThrArgAsnProProPro 102  
Db 247 AGAGAAATGGAGAAACAATGGTGGCAATGATCTCCACGCTCACCGGCTCTCTCTCT 306  
Qy 103 SerValProGlyProProProAsnProSerArgGluGlyGlySerProArg-----Pro 120  
Db 307 TCT-----CCTCTCTTAGGAGTAATGGAGATAATGGTGTAGCAGATCATCGCCA 357  
Qy 121 ProSer-----SerProSerProProSerPro-----SerSerAsp 132  
Db 358 CCAGGACACATGGAGGCTCTCGCTCAGACAACCTCTCTCTAGCGGAGGAGCAGTGA 417  
Qy 133 Gly-----LeuSerThrGlyValValValValGlyIleAlaIleGly 145  
Db 418 GGAGTGGAGGTGGAGAAGTAATACGAATACCGCATATAGTTGGTATTAGTTCGGA 477  
Qy 146 GlyValAlaLeuLeuValIleValThrLeuIleCysLeuLeuCysLysLysLysLysArgArg 165  
Db 478 GCTGGACTTTTGATGATCGTCTTATTATTGTGTCTTAGACGCAAAAGAGAGAGAAA 537  
Qy 166 ArgAspGluGluAspAlaTyrTyrValProProProProProGlyProLysAlaGly 185  
Db 538 -----GACTCCTCTCTAC-----CCTGAACCCATGAAGA 567  
Qy 186 GlyProTyrGlyGlyGlnGlnGlnTyrArgGlnAsnAlaThrProSerSerAsp 205  
Db 568 AACCAATAT-----CAATACTA-TGGAAACAACAACAACAACAATGCTTCACAGAA 617  
Qy 206 HisValValThrSerLeuProProProLysAlaProSerPro----- 221  
Db 618 -----TTATCCGAATTGGCACCTAAATTACAGAGGCCAAACCAACAATCTACTGGTGG 671  
Qy 222 -----ArgGln-ProProProProProPheMetSerSerSerGly----- 237  
Db 672 TTGGGAGGCGGTGGACCATCATCCGCTCTCTCTCCGCGATGCTTACACAGCGGAGAGA 731  
Qy 238 -GlySerAspTyrSer-----AspArgProValLeuProProProProGlyLeuVa 255  
Db 732 TTCTTCATGTACTACGGCCCATCACGCCAGTTTACCTCTCTCTCGCTCTCTAGC 791  
Qy 255 lLeuGlyPheSerLysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPh 275  
Db 792 CCTCGATTCAACAAGAGACATTTTACTTACCAAGAGCTTGGCGCTGCAACAGAGGGGTT 851  
Qy 275 eSerGluAlaAsnLeuLeuGlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuPr 295  
Db 852 TACGGATGCTAAACCTTTTGGACAGGAGGATTGGGTATGTCTCCATAAAGAGGCTTTGCC 911  
Qy 295 oSerGlyLysGluValAlaValLysGlnLeuLysValGlySerGlyGlnGlyGluAtgGl 315  
Db 912 TAGCGGGAAGAAGTAGCAGTTTAAGAGTTTAAAGCGGGTAGCCGACNAGGAGAGAGGGA 971  
Qy 315 uPheGlnAlaGluValGluIleIleSerArgValHisHisArgHisLeuValSerLeuVa 335  
Db 972 GTTTCAGCTGAGGTGATATCATTAGCCGTGTGCATCATCGTATCTTTCTTTCTTGGT 1031  
Qy 335 lGlyTyrCysIleAlaGlyAlaLysArgLeuLeuValTyrGluPheValProAsnAsnAs 355  
Db 1032 TGGATATTGCATAGCTATGCAGAGAGATGTTGGTTTATGAGTTTGTCTTCTTAAACAAAC 1091  
Qy 355 nLeuGluLeuHisLeuHisGlyGlyGlyArgProThrMetGluTrpSerThrArgLeuLy 375



```
QY 255 lLeuGlyPheSerIysSerThrPheThrTyrGluGluLeuAlaArgAlaThrAsnGlyPhe 275
Db 809 CCTCGGATTCACCAAGAGCACTTTTACTTACCAAGAGCTTCGGCTGCAACAGAGGGTT 868
QY 275 eSerGluAlaAsnLeuLeuGlyGlnGlyPheGlyTyrValHisIysGlyValLeuPr 295
Db 869 TACGGATGCTNACCTTTTGGACAGGAGGATTTGGGTATGTCTCAATAAAGAGTCTTGCC 928
QY 295 oSerGlyLeuGluValAlaValIysGlnLeuLeuValIglySerGlyGlnGlyGluArgG1 315
Db 929 TAGCGGAAAGAGTAGCAGTTAAGAGTTTAAAGCGGGTAGCGGACAAAGAGAGAGGGA 988
QY 315 uPheGluAlaGluValGluLeuLeuSerArgValHisHisHisHisLeuValSerLeuVa 335
Db 989 GTTTCAGCTGAGTGCATATCATTTAGCCGTGTGCATCATCGGTATCTTGTCTTTCTGGT 1048
QY 335 lGlyTyrCysAlaGlyAlaGlyAlaIysGlnLeuLeuValTyrGluPheValProAsnAsnAs 355
Db 1049 TGGATTTGATGCTGATGACAGAGATGTGGTTTATGAGTTTGTCTTCAACAAAC 1108
QY 355 nLeuGluLeuHisLeuHisGlyGluGlyArgProThrMetGluTrpSerThrArgLeuLy 375
Db 1109 TTTGGAATATCATCTTCATCGGAAATCTTCCGGTAATGGAGTCTCCACTAGGTTCGG 1168
QY 375 sIleAlaLeuGlySerAlaIysGlyLeuSerTyrLeuHisGluAspCysAsnProIysI1 395
Db 1169 TATCGCTTAGTGTCTCGAAAGGACTCGCTTACCTTCACGAAAGACTGCCATCTCGGAT 1228
QY 395 eIleHisArgAspIleIysAlaSerAsnIleLeuIleAspPheIysPheGluAlaIysVa 415
Db 1229 CATTCACCGGACATCAAGTCTCGCAATATCTCTTGACCTTCACTTATGCTATGGT 1288
QY 415 lAlaAspPheGlyLeuAlaIysIleAlaSerAspThrAsnThrHisValSerThrArgVa 435
Db 1289 GGCTGATTTTGGATTAGCTAAGTTAAACATCTGTATAACACACTCATCTACTCTGCTGT 1348
QY 435 lMetGlyThrPheGlyTyrLeuAlaProGluTyrAlaAlaSerGlyIysLeuThrGluLy 455
Db 1349 GATGGGAACCTTCGGATATCTAGCTCCAGAAATATGCTTCAAGCGGTAAATTAACCGGAA 1408
QY 455 sSerAspValPheSerPheGlyValValLeuLeuLeuLeuIleThrGlyArgArgProVa 475
Db 1409 ATCGGATGTTTCTTACGGAGTTATGTTATGGAACTTATTAACCTGAAACGACCGGT 1468
QY 475 lAspAlaAsnAsnValTyrValAspSerLeuValAspTrpAlaArgProLeuLeuAs 495
Db 1469 TGAT---AATAGCATCACCATGGACGACACCTTAGTAGATTGGGCTCGGCTCTTATGGC 1525
QY 495 nArgAlaSerGluGlnGlyAspPheGluGlyLeuAlaAspAlaIysMetAsnAsnGlyTy 515
Db 1526 TCGCGCGCTAGAAGATGGAACTTTAATAGAGCTCGCAGATCGCAGGGCTTGAAGCGCACTA 1585
QY 515 rAspArgGluGluMetAlaArgMetValAlaCysAlaAlaCysValArgHisSerAl 535
Db 1586 CAACCCGACAGAATGCTCGAATGGTACTTGTGGCCGCTGTAGCAATTCTGATTCGGG 1645
QY 535 aArgArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSe 555
Db 1646 GCCTAAACGTCCTAAAGATGACGAGATAGTAGAGCGCTTAGAGGAGAGTGTCTTAGA 1705
QY 555 rAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyrGlyGlySe 575
Db 1706 TGCTTTAAACGAAGGTGTGAAGCCAGGACACAGTAACGTTTACGGGTCAVTTGGGAGCAAG 1765
QY 575 rThrAspTyrAspSerSerGlnTyrAsnGluAspMetLysIysPheArgIysMetAlaLe 595
Db 1766 CTGGATTATAGTCAGACATCTTACAAATGCAGACATGAAGAAATTCAGACAGATAGCTTT 1825
QY 595 uGlyThrGlnGluTyr-----AsnAlaThrGlyGluTyr-SerAsnProThrSerAs 612
Db 1826 GTCGAGCCAAAGAAATTCACAGTCACTGCTGAGGAACATCTAGTAAATGATTCATAGAGA 1885
```

```
QY 612 pTyrGly 614
Db 1886 TATGGGA 1892
```

## RESULT 9

```
US-10-425-114-9117
; Sequence 9117, Application US/10425114
; Publication No. US20040034886A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9117
; LENGTH: 1424
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700833950_FLI
US-10-425-114-9117
```

```
Alignment Scores:
Pred. No.: 2,31e-104 Length: 1424
Score: 1737.50 Matches: 331
Percent Similarity: 88.73% Conservative: 47
Best Local Similarity: 77.70% Mismatches: 41
Query Match: 50.32% Indels: 7
DB: 13 Gaps: 4
```

US-10-086-464-2 (1-647) x US-10-425-114-9117 (1-1424)

```
QY 226 ProProProProProPheMetSerSerSerGly-----GlySerAspTyrSerAsp 243
Db 12 CCCCCACCCCGCTCTCTTCATCAGCAGCAGTGGGATCTGGATCAACAAATTCAGGC 71
QY 244 ArgProValLeuProProProSerProGlyLeuValLeuGlyPheSerIysSerThrPhe 263
Db 72 GGTGAATTTCTCTCTCTCTCTCAGGAATTCATTTGGGTTCTCTAAGAGCACATTC 131
QY 264 ThrTyrGluGluLeuAlaArgAlaThrAsnGlyPheSerGluAlaAsnLeuLeuGlyGln 283
Db 132 ACGTATGAGGAGTTGGCAGCGCACTGATGGCTTCTCTGATGCCAACCTCTCTGGACAA 191
QY 284 GlyGlyPheGlyTyrValHisIysGlyValLeuProSerGlyIysGluValAlaValIys 303
Db 192 GGAGGATTTGGATATGTGCACAGAGAAATTTCTCCCAACGCAAGGAGGTGGCAGTGAAG 251
QY 304 GlnLeuIysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGluIleIle 323
Db 252 CAATTGAAGGCTGGAAGCGGCAAGGGAGCGTGAATTCCAAGCTGAAGTTGAGATAATT 311
QY 324 SerArgValHisHisHisHisLeuValSerLeuValGlyTyrCysIleAlaGlyAlaIys 343
Db 312 AGCCGTGTCATCACAGCATCTGTGTTCTTTGGTGATGCTGCTGCTGCTGCTGCTGCTGCT 371
QY 344 ArgLeuLeuValTyrGluPheValProAsnAsnAsnLeuGluLeuHisLeuHisGlyGlu 363
Db 372 AGCTGCTGTTTATGAATTTGTTCCCAACACACATTCGAATTCATTTGCATGGAGAAA 431
QY 364 GlyArgProThrMetGluTrpSerThrArgLeuIysIleAlaLeuGlySerAlaIysGly 383
Db 432 GGAGAGACCTACCATGATTTGGCCCAAGACTAAGAAATTCCTTTAGGATCTGCTAAGGGA 491
QY 384 LeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLysAlaSer 403
```

Db 492 CTGGCGTATCTTCATGAAGATTGTCTATCTTAAGATCATCCATCGTGATATCAAAATCTGCC 551  
Qy 404 AsnIleLeuIleAapPheLysPheGluAlaLysValAlaAapPheGlyLeuAlaLysIle 423  
Db 552 AACATCTCTCGGATTTTAAAGTTTGAACAAAGGTTGCGAGATTTTGGTCTTGCAGAGTTT 611  
Qy 424 AlaSerAapThrAenThrHisValSerThrArgValMetGlyThrPheGlyTyrLeuAla 443  
Db 612 TCTTCTGATGCAATACCCATGTTTCTACTCGAGTGATGGGACTTTTGGGTATTGGCT 671  
Qy 444 ProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAapValPheSerPheGlyVal 463  
Db 672 CCAGAAATATGCTTCTAGTGGAAATCTCACAGAAATCAGATGTTTCTCTATAGGAATC 731  
Qy 464 ValLeuLeuGluLeuIleThrGlyArgArgProValAapAlaAenAasnValTyrValAap 483  
Db 732 ATGCTCTTGAGTTAATTAACCGGACGCGGCGAGTGTGATAAATCAAACTTACATGGAG 791  
Qy 484 AapSerLeuValAapTTPAlaAapProLeuLeuAenArgAlaSerGluGlnGlyAapPhe 503  
Db 792 GATAGTTTGGTAGATTGGGCTTAGGCTTTGCTCACAGAGCTTTGGAAGAGGATGATTTT 851  
Qy 504 GluGlyLeuAlaAapAlaLysMetAasnGlyTyrAapArgGluGluMetAlaArgMet 523  
Db 852 GATTCTATTATGACCCCAAGGCTCCAGAAATGACTATGATCTCTCATGAGATGGCAGATG 911  
Qy 524 ValAlaCysAlaAlaCysValArgHisSerAlaArgArgProAatqMetSerGln 543  
Db 912 GTGGCTTCTGCTGGGCTTGTATGCTATTCGTCGCAAGCGTTCACCAAGATGAGCGAG 971  
Qy 544 IleValArgAlaLeuGluGlyAasnValSerLeuSerAapLeuAenGluGlyMetArgPro 563  
Db 972 GTTGTCCGCGCTCTCGAAGGAGATGCTCTCTAGCAGATCTGAATGAAGGAATAAGACCT 1031  
Qy 564 GlyGlnSerAenValTyrSerSerTyrGlyGlySerThrAapTyrAapSerSerGlnTyr 583  
Db 1032 GGACACAGCATATGTATGTATCTCAT---GAAAGCTCAGATTTATGACACTGCACAGTAC 1088  
Qy 584 AasnGluAapMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyrAasnAlaThr 603  
Db 1089 AAGGAAGACATGAAGAAAGTTGAGGAATGCGATTTGGAACTCAGGAGTATGGTGCAAGC 1148  
Qy 604 GlyGluTyrSerAenProThrSerAapTyrGlyLeuTyrProSerGlySerSerSerGlu 623  
Db 1149 AGTGAGTACAGTCGCGCTACAGTGAGTGTGTTTAAACCCATCAGGCTCAAGTAGTAA 1208  
Qy 624 Gly-----GlnThrThrArgGluMetGluMetGlyLysIleLysArgThrGlyGln 640  
Db 1209 GCACAGAGCCGCCAAACCAAGGGAATGGAAATGAGAAAGATGAAG---AACAAATCAA 1265  
Qy 641 GlyTyrSerGlyProSer 646  
Db 1266 GGTTCAGTGAAGTTCT 1283

RESULT 10  
US-10-086-464-13  
; Sequence 13, Application US/10086464  
; Publication No. US2002019218A1  
; GENERAL INFORMATION:  
; APPLICANT: GORING, Daphne R. et al.  
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES  
; FILE REFERENCE: P 25,762-A USA  
; CURRENT APPLICATION NUMBER: US/10/086,464  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 10/069,304  
; PRIOR FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: PCT/CA00/00966  
; PRIOR FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/149,466  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: US 60/159,122  
; PRIOR FILING DATE: 1999-10-13  
; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 2025  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2025)  
US-10-086-464-13

Alignment Scores:  
Pred. No.: 6,22e-101 Length: 2025  
Score: 1687.00 Matches: 368  
Percent Similarity: 63.99% Conservative: 87  
Best Local Similarity: 51.76% Mismatches: 142  
Query Match: 48.86% Indels: 114  
Gaps: 25

US-10-086-464-2 (1-647) x US-10-086-464-13 (1-2025)

Qy 1 MetSerAlaProSerProGlyThrGlySerProSerProSerProSerAsnSerThr 20  
Db 16 GTGGATTTCATCTCTGCGCTTGAACCC---TCAATGGGACACCCCGTCAACGGGAACA 72  
Qy 21 ThrThrThrProProAlaSerAlaProProProThrThrProSerSerProProPro 40  
Db 73 TCG-----CGTCTAATGAGTCATCGCGGCAACACCCACCTTCTTCAACCAACCA 123  
Qy 41 ProSerThrIleProThrProProSerProSerSerArgSerThrProSerAlaProPro 60  
Db 124 TCATCA-----ATATCTGCTCTCCGCCAGATATCTCCGCTTCTTTTTCACCGCCG 174  
Qy 61 ProSerProProThr-----ProSerThrProGlySerProProProLeu 75  
Db 175 CTGCACCAACCAACGCAAGAAACGTCACCTCTCATCTCGTCTCATCGCCGCTGTT 234  
Qy 76 -----ProGln-----ProSerProAlaPro-----Thr 84  
Db 235 GTAGCTAATCGTCACCGCAGACTCCAGAGAACTCTTCTCCACCTGCACCTGAAGGCTCA 294  
Qy 85 ThrProGlySerProProAla---ProValThrPro-----ProThrArgAsn 99  
Db 295 ACTCTGTAAAGCCACCTGCACCAACCAACCAACCGTCGAACCAATCACCGGAAGACCA 354  
Qy 100 ProProPro-----Pro-----Pro-----Pro----- 102  
Db 355 ACTCTCTCTCTCTCTGTGTCATGATGACCGAAACAGAACCAATGCGGGAAACCAACAC 414  
Qy 103 -----SerValProGlyProProSerAsnProSerArg-----GluGlyGly 116  
Db 415 AGAGACGGCTCCACACCATCACCGCTCTGTCAGGGAACAGAACTTCCGGTGACGGTGGC 474  
Qy 117 SerProArgProProSerProSerProProSerProSerProSerProSerProSerThr 136  
Db 475 TCACCTTTCACCACTCGGTGCAAGCCCTCTCAGAAATAGTGAGATTCAGACTCATCA 534  
Qy 137 GlyValValValGlyIleAlaIleGlyGlyValAlaLeuValIleValThrLeuIle 156  
Db 535 TCG-----GGGCTT-----TTGCTTCTACTTGCAGTGTGATT 567  
Qy 157 CysLeuLeuCysLysLysLysArgArgArgAspGluAspAla-----Tyr 172  
Db 568 TGCATCTGTTGCAACAGGAAGAAAGAAATCTCTCAGGTCAACACCATGCACCTAC 627  
Qy 173 TyrValProProProProProProGlyProLysAla---GlyGlyProTyrGlyGln 191  
Db 628 TACAATAACAATCCTTATGAGGAGGACCCCTCAGGTAATGCTGTATTATCAAGGGA--- 684  
Qy 192 GlnGlnGlnTrpArgGlnGlnAsnAlaThrProSerSerAphisValValThrSerLeu 211  
Db 685 -----ACACCTCAAGATCATGTGGTG---AATATG 711  
Qy 212 ProProProProLysAlaProSerProProGlnProProProProProProPro 231







Dbb	1949	AAGAACTACGATAGANAATGAATAATGTTTCGGATGATCGAGGCTGCCGCACCCCTGTGTACGC	2008
Qy	533	HISserAlaArgArgProArgMetSerGlnIleValArgAlaLeuGluGlyAsnVal         : : : : :         : : : : :	552
Dbb	2009	CACTCATCGGTGAAGAGACCACGCATGAGTCAGGTGGTGAGAGCCTTAGATTCCCTTGGAT	2068
Qy	553	SerLeuSerAspLeuAsnGluGlyMetArgProGlyGlnSerAsnValTyrSerSerTyr : : :                 : : : : : : : :                 : : : : :	572
Dbb	2069	GAGTTTACGGATCTCAATAACGGAATGAACCGGGACAGAGTTCGGTG-----	2116
Qy	573	GlyGlySerThrAppTyrAppSerSerGlnTyrAsnGluAspMetLysLysPheArgLys : : :                 : : : : : : : :                 : : : : :	592
Dbb	2117	-----TTGATTCGGCGCAGCAATCTGCAAAATCAGAATGTTTTAGGAGG	2161
Qy	593	MetaLeuLeuGlyThrGlnGlnTyrAsnAlaThrGlyGluTyrSerAsnProThrSerAsp                 : : : : :                 : : : : :	612
Dbb	2162	ATGCTTTTGGAGGCCAAGAT-----AGTTCGGGTTCCTTCAGTAGCTCAGAGTAGC	2215
Qy	613	Tyr-GlyLeuTyrProSerGlySerSerSerGluGlyGlnThrArgGluMetGluMe : : :                 : : : : : : : :                 : : : : :	632
Dbb	2216	TGGAGGAGTA-----GAGATCACGACCCCAACAACCTGTCTTC	2251
Qy	632	tGlyLysIleLysArgThrGlyGlnGly	641
Dbb	2252	TCCCAAAATAAAAA---CTGGGCATCGGA	2276

## RESULT 13

```

US-10-425-114-7339
; Sequence 7339, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7339
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700652889_FLI
US-10-425-114-7339

```

Alignment Scores:		
Pred. No.:	3,378-91	Length:
Score:	1539.00	Matches:
Percent Similarity:	45.11%	Conservative:
Best Local Similarity:	53.11%	Mismatches:
Query Match:	44.57%	Indels:
DB:	13	Gaps:
		19
		2655

US-10-086-464-2 (1-647) x US-10-425-114-7339 (1-2655)

Qy	2	SerSerAlaProSerProGlyThrylSerProProSerProProSerAanSer-----	19
Db	98	TCATCTTCTCCAACA---AACACTTCCACCCACCATCACCTCCAGTTCTTCTCAGCCT	154
Qy	20	-----ThrThrThrThrProProProAlaSerAlaPro-----ProThrThr	34
Db	155	AATCAAAACAACAACAACAACACCAATTCCCCTGCCTTCATCTCTCTTGAC	214
Qy	35	ProSerSerProPro-----ProProSerThrIleProThrSerProProSerSer	52
Db	215	CCTTTGTGCTCCACCTCTTCTCCTCCTCAGGCAGTGCATTAACCCCTCTCTCTCATATA	274





```
Qy 243 AspArgProVal---LeuProProSerProGlyLeuValLeuGlyPheSerLysSer 261
Db 143 ACGGACCGCCCGTTCCTCCCTCCCTCACCACAACTTGCTGGGCTCAAAAGGGGC 202
Qy 262 ThrPheThrTyrGluGluLeuAlaArgAlaThrAngGlyPheSerGluAlaAsnLeu 281
Db 203 ACCTTCACTTACGAGGAATTAGCAGCTCCACCAACGGAATTCATGATGCAAAATTTGATA 262
Qy 282 GlyGlnGlyGlyPheGlyTyrValHisLysGlyValLeuProSerGlyLysGluValAla 301
Db 263 GGACAGGTGATTCGGCTATGTCCATAAGGGTGTGTGCTAGTGGAAAGAGTGGCA 322
Qy 302 ValLysGlnLeuLysValGlySerGlyGlnGlyGluArgGluPheGlnAlaGluValGlu 321
Db 323 GTTAAGAGTCTTAAAGCAGGTAGTGGCCAGAGAGCGAGAGTTCCAAGCTGAGATTGAC 382
Qy 322 IleIleSerArgValHisArgHisLeuValSerLeuValGlyTyrCysIleAlaGly 341
Db 383 ATCATTAGCCCGCTCCATCATCGCCATCTCTGTGCATCTGTGGATATCAATTTCTGGT 442
Qy 342 AlalyArgLeuLeuValTyrGluPheValProAsnAsnLeuGluLeuHisLeuHis 361
Db 443 GGCCAGAGATGTGGTCTATGATTTATTTCCCAATACACATTTGGAATATCACCTCCAT 502
Qy 362 GlyGluGlyArgProThrMetGluTrpSerThrArgLeuLysIleAlaLeuGlySerAla 381
Db 503 GGAAGGGTAGGCCCTACCATGGCACTAGCAATGCGGATTCGAATAGGATCCGCT 562
Qy 382 LysGlyLeuSerTyrLeuHisGluAspCysAsnProLysIleIleHisArgAspIleLys 401
Db 563 AAAGGGCTTGTTATCTTCATGAGACTGTCTATCTCGTATTATTCATTCGCGATCAAA 622
Qy 402 AlaSerAsnIleLeuIleAspPheLysPheGluAlaLysValAlaAspPheGlyLeuAla 421
Db 623 GCTGCAATGTCCTTATGATGATAGCTTCGAGCAAGGTGTGCTATTTGGATGGCT 682
Qy 422 LysIleAlaSerAspThrAsnThrHisValSerThrArgValMetGlyThrPheGlyTyr 441
Db 683 AAGTTGACTACTGATAATAATACTCATGTATCGACTCGTGTCTGGGAACTTTCGGGTAC 742
Qy 442 LeuAlaProGluTyrAlaAlaSerGlyLysLeuThrGluLysSerAspValPheSerPhe 461
Db 743 CTAGCCCTCGAATATGCATCAAGTGAAGAAATGACAGAGAAGTCTGACGTTTCTCATTT 802
Qy 462 GlyValValLeuLeuGluLeuIleThrGlyArgArgProValAspAlaAsnValTyr 481
Db 803 GGAGTCATGCTATTGGAACTCATAACTGGGAAGCGACCTGTTGATCACACAATGCC--- 859
Qy 482 ValAspAspSerLeuValAspTrpAlaArgProLeuLeuAsnArgAla---SerGluGln 500
Db 860 ATGGACGACAGCTAGTAGACTGGGCTCGACCACTTCTGACTCGTGGACTAGAGAGGAT 919
Qy 501 GlyAspPheGluGlyLeuAlaAspAlaLysMetAsnAsnGlyTyrAspArgGluGluMet 520
Db 920 GGCAACTTTGAGAGTTCGGTGTGATTCATTTTAGAGGGGAACTACGATCCTCAAGAACTG 979
Qy 521 AlaArgMetValAlaCysAlaAlaCysValArgHisSerAlaArgArgArgProArg 540
Db 980 TCAGAATGGCAGCTGTGTGCTCGCGGTAGCATTCGTCACTTCGCCAAAAAACGTCACAAA 1039
Qy 541 MetSerGlnIleValArgAlaLeuGluGlyAsnValSerLeuSerAspLeuAsnGluGly 560
Db 1040 ATGAGCCAGATTGTAAGAATATTGGAGAGAGATGTCTCACTGGATGACTTGAAGAATGGG 1099
Qy 561 MetArgProGlyGlnSerAsnValTyrSerSerTyrGlySerThrAspTyrAspSer 580
Db 1100 ATTAAGCCGGGCAAAATGTTGCTTTACAACTCTTCATCTAGCTCAGATCAGTATGACACA 1159
Qy 581 SerGlnTyrAsnGluAspMetLysLysPheArgLysMetAlaLeuGlyThrGlnGluTyr 600
Db 1160 ATGCAATACATGCTGATATGCAGAGTTCAGAAAGCGCAGTGTTCCTTAATAGCGAGAA 1219
```

```
Qy 601 AsnAlaThrGlyGluTyrSerAsnProThrSerAspTyrGlyLeuTyrProSerGlySer 620
Db 1220 TTTGGCACCAGC-----AGTGGCTCC 1240
Qy 621 SerSerGlu 623
Db 1241 AGTGGTGA 1249
```

Search completed: May 13, 2004, 03:06:24  
Job time : 728 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 14:53:24 ; Search time 7358 Seconds  
(without alignments)  
11451.326 Million cell updates/sec

Title: US-10-086-464-1  
Perfect score: 1944  
Sequence: 1 atgtcctcggcgccgtctcc.....atagtggacctctctcttaa 1944

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1944	100.0	1944	6	AX088876	AX088876 Sequence
2	1944	100.0	1944	6	AX825703	AX825703 Sequence
3	1944	100.0	2189	6	AX825705	AX825705 Sequence
4	1944	100.0	2189	8	AY028699	AY028699 Brassica
5	66	3.4	1959	6	AX825738	AX825738 Sequence
6	66	3.4	1959	8	BT008400	BT008400 Arabidops
7	66	3.4	1959	8	BT008409	BT008409 Arabidops
8	66	3.4	2098	8	AY128792	AY128792 Arabidops
9	66	3.4	2116	8	AY056788	AY056788 Arabidops
10	66	3.4	2188	8	AY059901	AY059901 Arabidops
11	66	3.4	2190	8	AY093065	AY093065 Arabidops
12	66	3.4	2257	8	AF370509	AF370509 Arabidops
13	51	2.6	79706	8	AB020746	AB020746 Arabidops
14	47	2.4	2324	8	AY089024	AY089024 Arabidops
15	29	1.5	80393	8	AP000382	AP000382 Arabidops
16	26	1.3	1515	6	AX825736	AX825736 Sequence
17	26	1.3	47428	6	AX059540	AX059540 Sequence
18	26	1.3	107826	2	AC124968	AC124968 Medicago
19	26	1.3	145453	8	AC012477	AC012477 Genomic S
20	26	1.3	183147	8	AC012392	AC012392 Genomic S
21	26	1.3	199987	8	ATCHRIV15	ATCHRIV15 Arabidops
22	24	1.2	898	8	AK109201	AK109201 Oryza sat
23	24	1.2	1389	6	AX654306	AX654306 Sequence
24	24	1.2	1922	8	AK070323	AK070323 Oryza sat
25	24	1.2	88450	8	AP004622	AP004622 Oryza sat
26	23	1.2	256	11	GI2716	GI2716 sWSS2029 Er
27	23	1.2	450	4	AF213398	AF213398 Ovis arie
28	23	1.2	897	6	AX546183	AX546183 Sequence
29	23	1.2	897	6	AX546184	AX546184 Sequence
30	23	1.2	897	6	AX546187	AX546187 Sequence
31	23	1.2	1185	8	AY224526	AY224526 Oryza sat
32	23	1.2	1281	6	AX546185	AX546185 Sequence
33	23	1.2	1620	9	HSU08818	HSU08818 Human activ
34	23	1.2	2016	9	HSU19348	HSU19348 Human (tpr-
35	23	1.2	3142	8	AY091180	AY091180 Arabidops
36	23	1.2	3407	8	AY064019	AY064019 Arabidops
37	23	1.2	3677	9	HSM808933	HSM808933 Homo sapi
38	23	1.2	4419	4	AB118945	AB118945 Canis fam
39	23	1.2	4586	9	HSMETPRO	HSMETPRO Homo sapien
40	23	1.2	4620	6	AX546174	AX546174 Sequence
41	23	1.2	4626	6	I96185	I96185 Sequence 22
42	23	1.2	4626	9	HUMMETPOA	HUMMETPOA Human MET p
43	23	1.2	4626	11	GI8239	GI8239 sWSS842 Eri
44	23	1.2	4823	4	AB112434	AB112434 Bos tauru
45	23	1.2	32173	9	AC004416	AC004416 Homo sapi

ALIGNMENTS

RESULT 1	AX088876	AX088876	1944 bp	DNA	linear	PAT 17-MAR-2001
LOCUS	Sequence 1 from Patent WO0114563.					
DEFINITION	AX088876					
ACCESSION	AX088876.1	GI:13397639				
VERSION	AX088876.1					
KEYWORDS	Brassica napus (rape)					
SOURCE	Brassica napus					
ORGANISM	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
	rosids; eurosids II; Brassicales; Brassicaceae; Brassica.					
REFERENCE	1					
AUTHORS	Goring,D. and Silva,N.					
TITLE	Proline-rich extensin-like receptor kinases					





```
QY 1261 GCTAAGATTGCTTCTGTATACAAACACGCACTGTATCAACACGCTGTGATGGGAACCTTTGGG 1320
Db 1261 GCTAAGATTGCTTCTGTATACAAACACGATGTATCAACACGCTGTGATGGGAACCTTTGGG 1320
QY 1321 TACTTGGCTCGGAATACGCTGCAAGCGGAAGAGCTCACGGAGAAGTCTGAGGTTTTCTCA 1380
Db 1321 TACTTGGCTCGGAATACGCTGCAAGCGGAAGAGCTCACGGAGAAGTCTGAGGTTTTCTCA 1380
QY 1381 TTGCGGTTGTGCTTTTGGAGCTCATTAATCTGGA CGTCGACCGCTTGATGCGCAACAATGTC 1440
Db 1381 TTGCGGTTGTGCTTTTGGAGCTCATTAATCTGGA CGTCGACCGCTTGATGCGCAACAATGTC 1440
QY 1441 TATGTAGATGACAGCTTAGTGTGACTGGCGACGACCATTTGCTTAACCGAGCATCTGAGCAA 1500
Db 1441 TATGTAGATGACAGCTTAGTGTGACTGGCGACGACCATTTGCTTAACCGAGCATCTGAGCAA 1500
QY 1501 GGAGACTTTGAGGGTTTAGCTGATGCAAGAGATGAATAATGGGTATGACAGAGAGGAGATG 1560
Db 1501 GGAGACTTTGAGGGTTTAGCTGATGCAAGAGATGAATAATGGGTATGACAGAGAGGAGATG 1560
QY 1561 GCTCGCATGTTGCTTGTGCTGCGGCTTGTGTTGCGCATTCAGCTCGCGCGACGCTCGC 1620
Db 1561 GCTCGCATGTTGCTTGTGCTGCGGCTTGTGTTGCGCATTCAGCTCGCGCGACGCTCGC 1620
QY 1621 ATGACCGAGATTGTCGTCGCTTAGAAGAAATGTATCACTGTCAAGATCTTAAACGAAGGG 1680
Db 1621 ATGACCGAGATTGTCGTCGCTTAGAAGAAATGTATCACTGTCAAGATCTTAAACGAAGGG 1680
QY 1681 ATGACACAGGTTCAAGCAATGTATACAGCTCATACGAGGAAGCACCGATTTATGACTCG 1740
Db 1681 ATGACACAGGTTCAAGCAATGTATACAGCTCATACGAGGAAGCACCGATTTATGACTCG 1740
QY 1741 AGCCAGTACAATGAAGACATGAAGAAATTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800
Db 1741 AGCCAGTACAATGAAGACATGAAGAAATTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800
QY 1801 AACGCCAGGTTGATACAGTAATCCGACCACTGACTATGACTGTACCCCTCTGGTTCA 1860
Db 1801 AACGCCAGGTTGATACAGTAATCCGACCACTGACTATGACTGTACCCCTCTGGTTCA 1860
QY 1861 AGCAGGAGGGGCAAAACACACGCAATGAGATGGGGAAGATTAAAGAAACCGGTGAG 1920
Db 1861 AGCAGGAGGGGCAAAACACACGCAATGAGATGGGGAAGATTAAAGAAACCGGTGAG 1920
QY 1921 GGTATAGTGACCTTCTCTTTAA 1944
Db 1921 GGTATAGTGACCTTCTCTTTAA 1944
```

```
RESULT 3
AX825705 LOCUS AX825705 2189 bp DNA linear PAT 11-DEC-2003
DEFINITION Sequence 3 from Patent WO03072763.
ACCESSION AX825705
VERSION AX825705.1 GI:39751232
KEYWORDS
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1
REFERENCE
AUTHORS Goring,D., Silva,N. and Haffani,Y.Z.
TITLE Increasing plant seed production
JOURNAL Patent: WO 03072763-A 3 04-SEP-2003;
Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)
FEATURES
location/Qualifiers
source
1..2189
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"
ORIGIN
```

```
Query Match 100.0%; Score 1944; DB 6; Length 2189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTCCTCGGCGCGCTCTCGGGGACTGGTTGCGCTCCATCTCCACCATCAAACTCCACA 60
Db 97 ATGTCCTCGGCGCGCTCTCGGGGACTGGTTGCGCTCCATCTCCACCATCAAACTCCACA 156
QY 61 ACACACATCTCTCTCAGAGCTTCGCTCTCTCTCCACACACACCTTCTCTCTCGCGCG 120
Db 157 ACCACCATCTCTCTCAGAGCTTCGCTCTCTCTCCACACACACCTTCTCTCTCGCGCG 216
QY 121 CCATCCACTATTCGAGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 217 CCATCCACTATTCGAGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276
QY 181 CCATCTCCACCACTCCATCTTACGCGGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCA 240
Db 277 CCATCTCCACCACTCCATCTTACGCGGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCA 336
QY 241 CCGGCTCAACTACGCGGGGATCTCCACCGGCACTGTCTTACTCTCTCTCTCTCTCTCTCT 300
Db 337 CCGGCTCAACTACGCGGGGATCTCCACCGGCACTGTCTTACTCTCTCTCTCTCTCTCTCT 396
QY 301 CCACCTTCAGTCCGAGGACCACTGTCCTCAATCCTTTCGCGGAAGGAGATCTCTCTCGACCT 360
Db 397 CCACCTTCAGTCCGAGGACCACTGTCCTCAATCCTTTCGCGGAAGGAGATCTCTCTCGACCT 456
QY 361 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Db 457 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 516
QY 421 GGAATCCGCAATCGGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Db 517 GGAATCCGCAATCGGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576
QY 481 AAGGAAGAAACGACGGAGAGACGAAAGAGATGCTTACTATGTTCTCTCTCTCTCTCTCTCT 540
Db 577 AAGGAAGAAACGACGGAGAGACGAAAGAGATGCTTACTATGTTCTCTCTCTCTCTCTCTCT 636
QY 541 GGTCCCAAGCGCGAGGACCTTACGGTGGACAGCAACAAATGGCGGCAACAAAACGCA 600
Db 637 GGTCCCAAGCGCGAGGACCTTACGGTGGACAGCAACAAATGGCGGCAACAAAACGCA 696
QY 601 ACACCACTGATCATGTCGTCGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db 697 ACACCACTGATCATGTCGTCGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 756
QY 661 CCACGGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Db 757 CCACGGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 816
QY 721 TACTCGGACCGTCCAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 817 TACTCGGACCGTCCAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 876
QY 781 AGCACTTTTCAATACGAGGAGCTAGTAGAGCCACCAATGGTTTCTCTCTCTCTCTCTCT 840
Db 877 AGCACTTTTCAATACGAGGAGCTAGTAGAGCCACCAATGGTTTCTCTCTCTCTCTCTCT 936
QY 841 TTAGGACAGCGGGTTGCGGTTACGTGACCAAAAGTGTGTTGCTCTCTCTCTCTCTCTCT 900
Db 937 TTAGGACAGCGGGTTGCGGTTACGTGACCAAAAGTGTGTTGCTCTCTCTCTCTCTCTCT 996
QY 901 GCTGTGAAGCAGTTGAAAGTTGGAGTGGTCAGGAGAGAGGAGATTTTCAGGAGAGAGTT 960
Db 997 GCTGTGAAGCAGTTGAAAGTTGGAGTGGTCAGGAGAGAGGAGATTTTCAGGAGAGAGTT 1056
QY 961 GAGATCATACGAGAGTTTCCACAGGCACTCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Db 1057 GAGATCATACGAGAGTTTCCACAGGCACTCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1116
```



QY 481 AAGAAGAAACGACGAGAGACGAAGAAGATGCTTACTATGTTCTCCGCGACCTCCTCCT 540  
Db 577 AAGAAGAAACGACGAGAGACGAAGAAGATGCTTACTATGTTCTCCGCGACCTCCTCCT 636  
QY 541 GGTCCCAAGCGGAGGACCTTACGGTGGACAGCAACAATGCGGCAACAAAACGCA 600  
Db 637 GGTCCCAAGCGGAGGACCTTACGGTGGACAGCAACAATGCGGCAACAAAACGCA 696  
QY 601 ACACCAACCGTCAGATCATGTGCTGACGTCACCTACCAACCAACCTTAAGGCTCCATCTCCA 660  
Db 697 ACACCAACCGTCAGATCATGTGCTGACGTCACCTACCAACCAACCTTAAGGCTCCATCTCCA 756  
QY 661 CCAAGGAAACCTCTCCACCTCCACCGGCTTCATGACGAGCAGCGGCGGCTCCGAC 720  
Db 757 CCAAGGAAACCTCTCCACCTCCACCGGCTTCATGACGAGCAGCGGCGGCTCCGAC 816  
QY 721 TACTCGGACCGTCAGATGTTCTCTCCACCGCTCCAGGCTTGTGTTAGGCTTCTCCAAA 780  
Db 817 TACTCGGACCGTCAGATGTTCTCTCCACCGCTCCAGGCTTGTGTTAGGCTTCTCCAAA 876  
QY 781 AGCACTTTTCAATACAGAGGAGTAGCTAGAGCCACCAATGGTTTCTCCGAGCGCAACTTG 840  
Db 877 AGCACTTTTCAATACAGAGGAGTAGCTAGAGCCACCAATGGTTTCTCCGAGCGCAACTTG 936  
QY 841 TTAGGACAAAGCGGCTTCGGTTACGTGCAAAAGGTGTTGCTAGTGGGAAAGAAGTT 900  
Db 937 TTAGGACAAAGCGGCTTCGGTTACGTGCAAAAGGTGTTGCTAGTGGGAAAGAAGTT 996  
QY 901 GCTGTGAAGCAGTTGAAAGTTGGGAGTGTGTCAGGAGAGAGGAGTTTCAGCGAGAGGTT 960  
Db 997 GCTGTGAAGCAGTTGAAAGTTGGGAGTGTGTCAGGAGAGAGGAGTTTCAGCGAGAGGTT 1056  
QY 961 GAGATCATCAGCAGAGTTTCAACACAGGATCTGTGTCTCTTGTGTTTATGTCATGCC 1020  
Db 1057 GAGATCATCAGCAGAGTTTCAACACAGGATCTGTGTCTCTTGTGTTTATGTCATGCC 1116  
QY 1021 GGTGCAAAAGATTCCTGTCTATAGATTTGTTCTCTAACAATCTCGAGCTTCACCTC 1080  
Db 1117 GGTGCAAAAGATTCCTGTCTATGATTTGTTCTCTAACAATCTCGAGCTTCACCTC 1176  
QY 1081 CATGCGAGGAGCGGCTTACATGAAATGGAGCACCAGATTGAAGATGCTCTTGGATCT 1140  
Db 1177 CATGCGAGGAGCGGCTTACATGAAATGGAGCACCAGATTGAAGATGCTCTTGGATCT 1236  
QY 1141 GCTAAGGACCTTCTTATCTTATCATGAAGATTGCAATCTCTAAATCAATTCACCGTGATC 1200  
Db 1237 GCTAAGGACCTTCTTATCTTATCATGAAGATTGCAATCTCTAAATCAATTCACCGTGATC 1296  
QY 1201 AAGGCTTCAACATATTGATAGATTTCAAGTTTGAAGCTAAGTTGCTGATTTGGTCTT 1260  
Db 1297 AAGGCTTCAACATATTGATAGATTTCAAGTTTGAAGCTAAGTTGCTGATTTGGTCTT 1356  
QY 1261 GCTAAGATTGCTTCTGATACAAACACGATGTATCAACACGTTGATCGGAACCTTTGGG 1320  
Db 1357 GCTAAGATTGCTTCTGATACAAACACGATGTATCAACACGTTGATCGGAACCTTTGGG 1416  
QY 1321 TACTTGGCTCCGGAATAGCTGCAAGCGGAAGCTCACGGAGAAGTCTGAAGTTTCTCA 1380  
Db 1417 TACTTGGCTCCGGAATAGCTGCAAGCGGAAGCTCACGGAGAAGTCTGAAGTTTCTCA 1476  
QY 1381 TTTGGCGTTGCTTTTGGAGCTCATTAATGGAAGTGTGATGCGCAATGTC 1440  
Db 1477 TTTGGCGTTGCTTTTGGAGCTCATTAATGGAAGTGTGATGCGCAATGTC 1536  
QY 1441 TATGTAGATGACGCTTATGTTGACCTGGGCAAGCAATGTTTAAACCGAGCATCTGAGCAA 1500  
Db 1537 TATGTAGATGACGCTTATGTTGACCTGGGCAAGCAATGTTTAAACCGAGCATCTGAGCAA 1596  
QY 1501 GGAGACTTTGAGGGTTTGTGATGCAAGAGATGAATATGGGTATGACAGAGAGGAGATG 1560  
Db 1597 GGAGACTTTGAGGGTTTGTGATGCAAGAGATGAATATGGGTATGACAGAGAGGAGATG 1656

QY 1561 GCTCGAGTGTGCTTGTGCTGCGGCTTGTGTTCCGCAATTCAGCTCGCGCAGACCTCGC 1620  
Db 1657 GCTCGAGTGTGCTTGTGCTGCGGCTTGTGTTCCGCAATTCAGCTCGCGCAGACCTCGC 1716  
QY 1621 ATGAGCCAGATTGTGCGTGTGAGAGGAAATGTATCACTGTGAGATCTTTAAACGAAGG 1680  
Db 1717 ATGAGCCAGATTGTGCGTGTGAGAGGAAATGTATCACTGTGAGATCTTTAAACGAAGG 1776  
QY 1681 ATGAGCCAGGTCAAGCAATGTATACAGCTCATACGAGGAGACCGATTATGACTCG 1740  
Db 1777 ATGAGCCAGGTCAAGCAATGTATACAGCTCATACGAGGAGACCGATTATGACTCG 1836  
QY 1741 AGCCAGTACATGAAGACATGAAGCTTTAGGAAAATGGCACTTGAACCTCAAGAGTAC 1800  
Db 1837 AGCCAGTACATGAAGACATGAAGCTTTAGGAAAATGGCACTTGAACCTCAAGAGTAC 1896  
QY 1801 AACCCACCGGTGAGTACAGTAATCCGACCAGTGACTATGAGCTGTACCCGTCTGGTTCA 1860  
Db 1897 AACCCACCGGTGAGTACAGTAATCCGACCAGTGACTATGAGCTGTACCCGTCTGGTTCA 1956  
QY 1861 AGCAGCGAGGCGCCAAACCAACGCGAAATGGAGATGGGGAAGATTAGAGAACCGGTGAG 1920  
Db 1957 AGCAGCGAGGCGCCAAACCAACGCGAAATGGAGATGGGGAAGATTAGAGAACCGGTGAG 2016  
QY 1921 GGTATATAGTGAACCTTCTCTTTAA 1944  
Db 2017 GGTATATAGTGAACCTTCTCTTTAA 2040

RESULT 5  
AX825738  
LOCUS AX825738  
DEFINITION Sequence 36 from Patent WO03072763.  
ACCESSION AX825738  
VERSION AX825738.1 GI:39751255  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.

REFERENCE 1  
AUTHORS Goring, D., Silva, N. and Haffani, Y. Z.  
TITLE Increasing plant seed production  
JOURNAL Patent: WO 03072763-A 36 04-SEP-2003;  
Goring, Daphne (CA) ; Silva, Nancy (CA) ; Haffani, Yosr, Z. (CA)  
FEATURES  
source 1. .1959  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Query Match 3.4%; Score 66; DB 6; Length 1959;  
Best Local Similarity 100.0%; Pred. No. 4.6e-24;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1432 AACAAATGCTATGTAGATGACGCTTAGTTGACTGGGCAAGCATTGCTTTAACCGAGCA 1491  
Db 1447 AACAAATGCTATGTAGATGACGCTTAGTTGACTGGGCAAGCATTGCTTTAACCGAGCA 1506  
QY 1492 TCTGAG 1497  
Db 1507 TCTGAG 1512

RESULT 6  
BT008400  
LOCUS BT008400  
DEFINITION Arabidopsis thaliana At3g24550 gene, complete cds.  
ACCESSION BT008400  
VERSION BT008400.1 GI:30725473  
KEYWORDS FLI\_CDNA.









```
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="III"
/clone="RAF11-07-N23"
/ecotype="Columbia"
/note="This clone is in pBluescript"
1..2257
/gene="MOB24.13"
107..2065
/gene="MOB24.13"
/codon_start=1
/product="protein kinase-like protein"
/protein_id="AAK3886.1"
/db_xref="GI:13877617"
/translation="MSTAPSPGTTTSPSPPTTSTTTTTPPAASPPPTTTPSSPP
PSPSTNTPSPSPSLPPLPSPSPGSLTPPIQPSAPITSPSPPTTSPNRPSP
PSPNQPTSGTPTPTNTKSPSPSDSLSTGVVGIAGVAILLITLCL
LCKRRRRHDEAAYVYPPPPSPGKAGPYGQQQWQOONASRPSDHHVTSLLP
PKPSPRPKPPPPAPFPMSSSGSDYSDLPVLPSPGLVLFKSTYEIYLSRA
TNGFSEANLLQGQGFYVHGKILPSGKEVAVKQKAGSGGQREFQAEVEIISRVHR
HLVSLIGVCMAGVORLLVYFVPPNNLEFHLHGKGRPTMEWSTLKIAGSAGLSYL
HEDCNPKLIHRDIKASNLIDFKFAKVADEGLAKIASDNTNTHVSTVWMTFGYLAPE
YAASGKLTKEDSFSGVYVLELITGRPVDDANNVYVDDSLVDWARPLNRASEEGDF
EGLADSKMGNEYDREEMARVACAAACVRHSARRPRMSQIVRALEGNVSLDNEGM
RPGHSNVYSSGSDYDTSQYNDMMIKFRKALGTQYGTGTEYSNPTSDYGLYPSG
SSSEQATREMKGKIKKTGGYSGPSL"

FEATURES             source
CDS
1..79706
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOB24"
/clone_lib="Mitsui P1"
2398..3399
/note="gene id:MOB24.1
unknown protein"
/evidence=not experimental
/protein_id="BAB01995.1"
/db_xref="GI:9294038"
/translation="MGDSDEGTGPKKFKYSLNQNHPMYSRPIKRAHYNEEDED
VEGEEDPOGGYIRGNRFQKROKPNKYVSGFEPAGSDAKVAYDREQAFVLLEW
GDRFLQKRLSRNEDWNEVAEKSEELMERKSETQRRMIDDLKRYREKIKVKS
GLVSKSWSPFNKLDMLLCKSLGACVDSGEFVFNNTKYLVKSNFGFDEMDSP
GDSEEDDEDVEYERKYNDAASYMLADSVERFGYKYEKMKSKKQMKLEKQR
ADFDLLELQKQIVDRAQSEIARLREEEENHGGDDDESEDEMDNDVNLSD"
complement(join(3794..4276,4363..4490,4577..4673,
4757..5179,5354..5548,5693..5970,6088..6214,6297..6496,
6601..6648,6743..6776,6882..7013,7464..7559,7677..7743,
8162..8217,8290..8385,8470..8998,9122..9462))
/note="gene id:MOB24.2"
/codon_start=1
/evidence=not experimental
/product="DNA repair protein MutS"
/protein_id="BAB01996.1"
/db_xref="GI:9294039"
/translation="MQQRSLTSFFQKTAATTGLVSGDAASGGSGGPRFNVKEG
DAKGDASVRFVSVKSVDEVRGTDTPEKVPRRVLPSPGKPAESAGDASSLFSNIMHKF
VKYDDRDGCSGERSRDVPLNDSSLCHKANDVIQPRSNNGKTQERNHAFSFGRAEL
RSVEDLGVDVPGPPTPGRRPRASRLKRVLEDMTFKEDKVPVLDSNKKLMLQDPV
CGEKEVNGFTKFEWLESRIIDANRRRDPDLYDRKTLHIPVDFVKMSAQKQYWS
VKSEYMDIVLFVKFYELYELDABELKELDKMTSGVGCROQVIGISEGIDEAV
KLLARGYKCRVTEQLESDQAKARGANTII PKLVQVLPSTASEGNIGPDVHLLA
IKELMELQKCVTVGFADFCAALRFWGSISDDASCAALGALLMQVSPKEVLYDSK
GLSREQAALRKTYLTGSTAVQLAPVPQWMDTDAAGVNIIESNGYFKGSSSSWNC
VGLNCECDVALSGALYGLNDNCVSTGTGLLRNWKCHPLKDVESINRKLDDVVEFTANS
NNSCDGSPGTGLYKLDNCSPTGKLLRNWIKHIDIPFYQVIRGLRIDGQVWNLIF
MIDTQYGLKHLPLDLRLGRKISSVRSASSVLPALGKVKLVKQVAFQVIRGFRSG
IQDILLALQKESNMMLLYLCKLPILVKGSGLELFLSQEAAIDSPFPNQNDVDE
NAETLTLLLEFIERATQWSEVHTTISCLDVLFSFAIATASLSAGSWARPVI FPESEAT
DQNKTKGILKIQLGWHHPFAVAADQLVPFNDILGEARRSGSIHPRSLITGPNM
GGKSTLLRATCLAVIFAQGCYVPCSEISLVDTITFRUGASDITGGSTLVECT
ETASVQNATQDSLVILDELRGTSTFDGYAIYSVFRHLVEKQVCRMLFQALMAGI
KEFASHPRTSKHMACAFKRSYDQPRGQDQDILVFLYRLTEGACPSYGIQVAMAGI
PNOVDTAGGAQAMKRSIGENFKSELSESPSSLEDMLKSLVIGISRVAHNAPIGE
DYDVTLCFLWHEITKSSYCVPK"
10131..10577
/note="gene id:MOB24.3"
/codon_start=1
/evidence=not experimental
/product="ethylene-responsive transcriptional
```

CDS  
coactivator-like protein"  
/protein\_id="BAB01997.1"  
/db\_xref="GI:9294040"  
/translation="MPERYPAQTQWPEVVLVHKSQKSDLPDKPAVNAALRNGVAV  
QTVKFDAGSKGKSTAVPINTKCLEETEPAAMDVKAELVRLMIQKARLEKMSQ  
ADLAKQINERTQVQVEYENGKAVNQAVLAKMEKVLGKLRGKIGK"  
complement(join(11108..11381,11467..11528,11699..11836,  
11934..12071,12161..12334,12421..12510,12718..13104,  
13623..13765,14306..14405))  
/note="gene\_id:MOB24.4"  
/codon\_start=1  
/evidence=not experimental  
/product="aldehyde dehydrogenase"  
/protein\_id="BAB01998.1"  
/db\_xref="GI:9294041"  
/translation="MENGKNGCATTVKLPEIKFTKLFINGQFIDAAGKTFETIDPRN  
GEVATTAEGKDEQVDVLAANAARVAFDHPGPRMTGFERAKLINKFADLAEENIEELA  
KLDVADGSKLPOLGKYADI PATAGHERVYAGAAADKI HGETUKMTRFSLFGYTLKEPIG  
VVGNIIPNPFSPIMPAKVAPAMAAGCTMVVKPAEQTSLSALFAHLKSKEAGIPDGVL  
NIVTGFSTAGAAIASHMDVDKVFSTGTVGRKIMQAAASNLKVSLELGGKSPILL  
IFNDADIDKAADLALLGCFYKNGKEICVASSRVFQVEGIYDKVVEKLEKAKDWTGDP  
FDSTARQGPVQDKQFEKILSYIEHGKNEGATLLTGGKAIGDKGYFIQPTIFADVTE  
MKIYQDEIFGVMSLMKFKTVEEGI KCANNTKYGLAAILSQDIDLINTVSRSIKAGI  
IWNCFYGFDDCDPVGGMKSGNCRSGMDALDNLVOTKSVVPLHNSPMW"  
complement(join(19742..19933,20400..20657))  
/note="gb|AAD25142.1  
gene\_id:MOB24.5  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB01999.1"  
/db\_xref="GI:9294042"  
/translation="WTPUSIASPSPWMLSPPLSSSLNRAAPISFPSPSRINHGF  
FSPKSFPLPPMPYDPPEEEEEKEEETDDPEKEDEQPDNRS"  
join(23402..23459,23819..24021)  
/note="gene\_id:MOB24.6  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02000.1"  
/db\_xref="GI:9294043"  
/translation="WKNASLKLPLLIFILVITSNLGAEARKLTGTADVIVVEGEAASPO  
YGIIDENTLNKSPCKRDICTQCRGRCNLIQKCECL"  
complement(join(25945..26640,26714..26824,26914..27012,  
27119..27224,27291..27360,27436..27583))  
/note="contains similarity to ubiquitin conjugating enzyme  
gene\_id:MOB24.7"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02001.1"  
/db\_xref="GI:9294044"  
/translation="MAQARLSRMQKELKLLSDPPPHGASFPHLSSAANGSGDFSTF  
STIDAIQIGPDTVYANGIFNLKIQIPERYFPQPIVSFATPIVHPNIDNSGRICLDI  
LNLPPKQAWQSLNISTVLTSMRLLESPNDDGLMCEVSRKYKNTQPTDYKAREMT  
EKYAKVADGSLQIKNIGDEKSGESGNVVKLTVESLSLAHTVKTETAGDEQ  
EDNGKRKAVVGFEGNSFGNDGIKRSNKLALPQSKDLCGBEELRGVSAAC  
KENKKPNLNGKLSLGLQPCNDTLASFRTSAKSDNNRLSKLSRSLPLGELNEVSK  
PEVLAQDTMKLEMNQEDARSURGEFENSVEETSMABSIIVVLDDSDSGQBEERVS  
SRSRLSLAKRVLKCRP"  
28636..30168  
/note="emb|CAB79678.1  
gene\_id:MOB24.8  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02002.1"  
/db\_xref="GI:9294045"  
/translation="MSSLIAYLSLPSNFSPLLKLSRILSNLWLLVTRSTNTKRKRWCL  
AIPIYHLHATPLWKSSTMLLRFSQMVPEPTPYETISLSRAFILOAMSGETLLMD  
TTSIFGLTRILWFFSSASPHGQNTSTFPSPENDASLSSPFYNSRRHHHHLSEGKL  
GLQLVGLVFLPLRYRMPISKDDPKLLPKRSLDNLVSLISFQSLSYHLIASIDRGHH  
TNHNLVLWLHKNLSRLTLPLVSSRTFTVTGICLKNGINGIMFPTPTSRWCDCIHPISATI

CDS  
CISPTHLGISGKQSSSRVKAHISRISKEVPNLFELIITPSPHALVVKGLKKIGIMI  
TSRPSGDYORFENLSHSYPLITREKOKRFNI IKFSDLOSGLKNGKIMISSPWSDDYR  
PSTRLDLDTPIVTHLLGDEQLQFLLKVSARTSALEPSTSRLLTITWLLSISDFVED  
SFNTRDUTCAIMPYSLWLKALMFMFSFHIIYTHILVLGNASFNVINLVIFESLLLIIV  
LI"  
join(32831..33058,33143..33907)  
/note="contains similarity to heat shock transcription  
factor  
gene\_id:MOB24.9"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02003.1"  
/db\_xref="GI:9294046"  
/translation="MEDDNSNNNNNVIAPFIVKTYQMVNDSTDLITWGPANNSF  
IVDPLDFSORIILPAYFKNNFSFVRLQNTYGFYKVDPPRWFEPANEHFLRGQKHLN  
NIARRHARGMGLDEGEIVREIERKEEQRLEAEIQRMNRRIBATRPRQPMMA  
FLYKVEDPDLPLPRMLLEKSTQQQQVSDKKRVRTMTVKSEEEVEDEGRVFRV  
MSSSTSPSSTENLYRNHSDGWIPTWMTQFGSVETGLVAKLSNSTSSSLTS  
TFSLPSVNGGGGGCGSGIOGERRYKETAATFGVGVESNPPTTPYPFSLFRGGP"  
complement(join(37054..37590,37670..37809,37894..38197,  
38288..38398,38486..38584,38662..38772,39019..39066,  
39311..39367))  
/note="gene\_id:MOB24.10  
unknown protein"  
/codon\_start=1  
Query Match 2.6%; Score 51; DB 8; Length 79706;  
Best Local Similarity 100.0%; Pred. No. 7.2e-16;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1579 GTCGCGCTGTGTTCGCCATTTCAGCTCGCGCAGACTCGCATGAGCCAG 1629  
|||||GCTCGCGCTGTGTTCGCCATTTCAGCTCGCGCAGACTCGCATGAGCCAG 54058  
Db 54008 GCTCGCGCTGTGTTCGCCATTTCAGCTCGCGCAGACTCGCATGAGCCAG 54058  
RESULT 14  
AV089024 2324 bp mRNA linear PLN 14-APR-2003  
LOCUS Arabidopsis thaliana clone 17909 mRNA, complete sequence.  
DEFINITION Arabidopsis thaliana (thale cress)  
ACCESSION AV089024  
VERSION AV089024.1 GI:21407798  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana (thale cress)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 2324)  
Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,  
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.  
Full-length messenger RNA sequences greatly improve genome  
annotation  
Genome Biol. 3 (6), RESEARCH0029 (2002)  
JOURNAL 22088475  
MEDLINE 12093376  
PUBMED 12093376  
REFERENCE 2 (bases 1 to 2324)  
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and  
Feldmann,K.  
Full-length cDNA from Arabidopsis thaliana  
Unpublished  
3 (bases 1 to 2324)  
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and  
Feldmann,K.  
Direct Submission  
Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,  
Malibu, CA 90265, USA  
This clone sequence is one of 5,000 Ceres full-length cDNAs made  
available to TIGR and Genbank. The following quality assessment of  
this set was done by comparison with known proteins: two percent of  
the clones are estimated to be 5'-truncated; less than one percent  
are 3'-truncated; approximately two percent represent alternative  
splice variants, including unspliced introns and spliced exons; one  
percent may contain premature stop codons; five percent may have

frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or Llaer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Genset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

## FEATURES

Location/Qualifiers  
1..2324  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="17909"

## ORIGIN

Query Match 2.4%; Score 47; DB 8; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 1.4e-13;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1595 GCTTGTGTCGCATTGACGTCGCCGACAGCTCGCATGACCCAGAT 1631  
|||||  
Db 1700 GCTTGTGTCGCATTGACGTCGCCGACAGCTCGCATGACCCAGAT 1746  
|||||

## RESULT 15

AP000382/c 80393 bp DNA linear PLN 27-DEC-2000  
LOCUS Arabidopsis thaliana genomic DNA, chromosome 3, TAC clone:K7M2.  
DEFINITION AP000382 BA000014  
ACCESSION AP000382.1 GI:5672513  
VERSION

## KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

## ORGANISM

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

## REFERENCE

AUTHORS Kaneko, T., Katoh, T., Sato, S., Nakamura, A., Asamizu, E. and Tabata, S.  
TITLE Structural analysis of Arabidopsis thaliana chromosome 3. II.  
Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC  
and BAC clones

DNA Res. 7 (3), 217-221 (2000)

## JOURNAL

MEDLINE 20363099

PUBMED 10907853

2 (bases 1 to 80393)

AUTHORS Kaneko, T., Katoh, T., Sato, S., Nakamura, Y., Asamizu, E. and Tabata, S.

TITLE Direct Submission

JOURNAL Submitted (28-JUL-1999) Yasukazu Nakamura, Karusa DNA Research

Institute, Department of Plant Gene Research; 1532-3, Yana,

Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,

Tel: 81-438-52-3935, Fax: 81-438-52-3934)

Address for correspondence: kaos@kazusa.or.jp

For the latest information on annotation of this clone, please see

Genes with similarity to proteins in the databases are described in

'product' or 'note' qualifiers. Genes that have no significant

protein similarity are described as 'unknown protein'.

The software programs used to predict genes include: Grail

(Informatics Group, Oak Ridge National Laboratory,

http://compbio.ornl.gov/Grail-1.3/),

GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),

NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of

Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and

SplicePredictor (Volker Brendel, Stanford University,

http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).

Genes encoding tRNAs are predicted by tRNAscan-SE

(Sean Eddy, Washington University School of Medicine, St. Louis,

http://genome.wustl.edu/eddy/tRNAscan-SE/).

This sequence may not be the entire insert of this clone. It may be

shorter because we remove overlaps between neighboring submissions.

The 5' clone is K13K6 and the 3' clone is MXP5.

## FEATURES

source  
Location/Qualifiers  
1..80393  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="K7M2"  
/clone\_lib="Mitsui TAC"  
complement(1077..2201)  
/note="gb|AA03494.1  
gene id:K7M2.2  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02925.1"  
/db\_xref="GI:11994687"  
/translation="MADADELMTKNGVKTSQSSVQTRLSRSVLRLLNPELTRELRRLK  
DNLNRKIDHDEKPEENPILNPKMIPRIIVYKPKDDVERKMKDLKQKLTLL  
KAJEBRLTRBELRVYTCCLKDFKSHRIDYQFVKSLQRLIDKYNLYQRFVDIATGHG  
EDAGTQVVDNDPQDTESDKIKLEGKKGMLPPKESDOPERLKKRKRTERATPNY  
KLIPKSEQTPVSVTLNVTWVNSYDIAQKNLTDIEKMYNWDQMFELMDLVGLFT  
SAAKNAEVSINGERDLKLGKFKYCAENLYGRMDLEIVKENHORVLPAILNRLNQKL  
REVTLVRELRKPVLKQTIKLSRQSGTAQESQVKTRIT"  
complement(join(3781..4009,4102..4277))  
/note="gene id:K7M2.3  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02926.1"  
/db\_xref="GI:11994688"  
/translation="MRMVQLLFKLIQVGTAVFELSNEIGNKNLLLIHHAGSLLS  
MFGNDCLIAAFDKKGLTLIASEFSDYFENLVKDRVGNVYVQRLINGFKTKTGIDLPHS  
LTSVLVTRSHLCKHRYQVIEAFDRSTRLA"  
complement(6378..6587)  
/note="gene id:K7M2.4  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02927.1"  
/db\_xref="GI:11994689"  
/translation="MRPMVYRDFLGMEYFGYSMPMQMEIDEDDFGPMQMDFEYGGIS  
PFQMKPDSDFPNKFEDDDFDSIN"  
complement(8640..10130)  
/note="gene id:K7M2.5"  
/codon\_start=1  
/evidence=not experimental  
/product="ammonium transporter"  
/protein\_id="BAB02928.1"  
/db\_xref="GI:11994690"  
/translation="MSGATCSAADIALLGPNATAADYICQGLGVNKNKFTDAAYA  
IDNTYLLFSAYLVFAMQLGAMLCAGSVRAKNTNMLTNVLDAAAGGLFYILFGYAF  
AFGSSDGTGRHFNGLQNPFTLTSDYSFELYQWFAIAAAGTSGSIAERTKFAVL  
IYSSFTGTFYVYVSHWSPDGMWSPFRSEDLFGTGAIDFAGSGVHVHVGGLW  
GALIEGPRIGRFDGGHATALRGHSASLVGLTFLWFGWYGNPQSFTKLIPYNSG  
SNYQSGSGIGRTAVTTLSGCTAALTFLFKRLLSGHWNVTDVNCGLGSPFAITAC  
SVYDWAIVCGFVASVLIGCNLAELKIDPLEAQHWYDQGGCAGGLIFVGLFAKE  
KYINEYVAGSPGRHYGLFMWGGKGLQAQLVQIIVIVGWSATMGTLTFLIKKLNILR  
ISQHEMRGMDLAGHGGFYIYHNDNDDDSIGVPGSPVPRAPNPPAV"  
complement(13098..14594)  
/note="gb|AAD54638.1  
gene\_id:K7M2.6"  
/codon\_start=1  
/evidence=not experimental  
/product="ammonium transporter"  
/protein\_id="BAB02929.1"  
/db\_xref="GI:11994691"  
/translation="MSGATCSAADIALLGPNATAADYICQGLGVNKNKFTDAAYA  
IDNTYLLFSAYLVFAMQLGAMLCAGSVRAKNTNMLTNVLDAAAGGLFYILFGYAF  
AFGSSDGTGRHFNGLQNPFTLTSDYSFELYQWFAIAAAGTSGSIAERTQFAVL  
IYSSFTGTFYVYVSHWSPDGMWSPFRSEDLFGTGAIDFAGSGVHVHVGGLW  
GALIEGPRIGRFEKGRALALRGHSASLVGLTFLWFGWYGNPQSFTKLIPYNSG  
GSNYQSGSGIGRTAVTTLSGCTAALTFLFKRLLSGHWNVTDVNCGLGSPFAITAC

## CDS

## CDS

## CDS

## CDS

## CDS

CSVBPWAAIYCGWASVVLGCKNAELVOYDDPLEAQLHGGCGAGWGLFVGLFAK  
EKYLNEVTGATPGPYGLFMCGGKLLGAQLVQLILVIVGWSATWGTILFFILKRLNLL  
RISQHEHQGMDMTRHGFAIYIYHNDDESHRVDPGSPFPNSATPPRV"  
complement (join (18576, .19104,19290, .19570))  
/note="gene\_id:K7M2.7"  
/codon\_start=1  
/evidence=not experimental  
/product="MYB-like DNA-binding domain protein"  
/protein\_id="BAB02930.1"  
/db\_xref="GI:11994692"

CDS

/translation="MSLWGGMGGMGVEGWRKGPWTAEDRLIIDYVQLHGRWN  
SVRLAGIKRWKSRCLRWNYLRLPDLKRGQITPHEETIILELHAKGNWSTIARSL  
PGRTDNEIKNYRTHFKKTKSPNSAEKTKNRILKQQEQOQORQLOOQLLOFN  
QIDMKXIMSLDDNNNGDNTFSSSGESGALYVPHQIHTHTTSCEPNSNGYYPV  
VFVTPEANVNDAIMDGLWNLDLFEQGSFGGAACAPRKHYPONWVPPC"  
complement (join (27839, .28076,28172, .28286,28388, .28424,  
28537, .28612,28792, .28847,28948, .29017,29396, .29796))  
/note="gene\_id:K7M2.8  
ref|NP\_001196.1  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/product="MYB-like DNA-binding domain protein"  
/protein\_id="BAB02930.1"  
/db\_xref="GI:11994692"

CDS

/translation="MSLWGGMGGMGVEGWRKGPWTAEDRLIIDYVQLHGRWN  
SVRLAGIKRWKSRCLRWNYLRLPDLKRGQITPHEETIILELHAKGNWSTIARSL  
PGRTDNEIKNYRTHFKKTKSPNSAEKTKNRILKQQEQOQORQLOOQLLOFN  
QIDMKXIMSLDDNNNGDNTFSSSGESGALYVPHQIHTHTTSCEPNSNGYYPV  
VFVTPEANVNDAIMDGLWNLDLFEQGSFGGAACAPRKHYPONWVPPC"  
complement (join (27839, .28076,28172, .28286,28388, .28424,  
28537, .28612,28792, .28847,28948, .29017,29396, .29796))  
/note="gene\_id:K7M2.8  
ref|NP\_001196.1  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02931.1"  
/db\_xref="GI:11994693"  
/translation="MPRELLADCTGRPSLFISLLITKLCFSFIIVEICSYDMDEVVVE  
VEKTRWEWEAYEKTIGIHSIQIGKRRGGGEEKPSLQRLNGLAQDGLSLNSLQ  
FNLLDLPQSPDQVQSTQSLLETQWKNQHSRLNLSANLQAKDNKNRKAQAEEREL  
LIGGTSTVLRKQANAGVTSDAESITSLRRSRQLMVOEVERSTNIVAFDESTG  
VLKXAESYKGRSLSLRNLSTMRQDVIDRIILIVGSLFVFAVYVVKRIGI  
LKLQRMATAIKAOLAGKAANGVDVMPGQDFGNTVPTVNIPLQORVHDEL"  
complement (join (30469, .30702,30798, .31445,31543, .31842,  
31969, .32026,32100, .32243,32335, .32591,32720, .32929,  
33121, .33294,33402, .33469,33563, .33621,33705, .33780,  
33877, .33982,34075, .34184,34450, .34614,34726, .34811,  
34942, .35006,35096, .35143,35598, .35717,35851, .35926,  
36187, .36233))  
/note="contains similarity to mismatch repair protein MutS  
gene\_id:K7M2.9"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB02932.1"  
/db\_xref="GI:11994694"

CDS

/translation="MYTNLLGLDPSLRNGLKGNLAWMLQFKSRPREVLLCRVGE  
FYAIGDACILVEAGLNPFGGLKSDSIPKAGCPTNNLQTLDDLTRNGYSCVIEE  
VQGPTPARSRRGRFISGHAHPGSPYVGLVGDHLDLPDPMPVVGISRSARGYCMIS  
IFETMKAYSLLDDGLTEALVTKLRTRCHLFLHASLRHNASGTRMGEGEGGLLWG  
ECSSRNFWFEGDTLSLLSRVQVYGLDDDEVFRNVNPKSRPRPLHLGTATQIGA  
LPTEGICLLVLLPSTCSGLPSLVVRLDLPAYDIALKIQETCKLMSTVTCIPE  
FTCVSSAKLVLLQREANYIEFCRIKNVLDLVLHMHAEILVELKLLMDPTVATG  
LKIDPDTVNECHNASOTIGEMISLDENESHQNVSKDNVNEFFYMESSWRGVKG  
IHIIEEITQVKEKSABALSLVAEDPHPIISRIKATTASLGGPGKEIAYAREHESVWFK  
GKRFTPSIWAGTAGDQIKQKLPALDSKGVGEWFTTPKVEIALVRYHEASENAKA  
RVLELLRSLVKLOTKIVLVFASMLVISKALFSHACERRRRKRWVFTPLVGLSLDEG  
AKPLDASRMKLTGLSPYFEDVSCTAVHNTVDMQSLFLHGPNGGGKSLRSICAA  
ALLGSLMVAESACIPEFDSIMLHMSYDSPVDGKSSQFVEMSEIRSIYQATRS  
LVLDIEICRGTEATAGTSVESLSTGCLGIVSTHLHGIFSLPLTAKNIYKAM  
GAENVGQGTPTKLTQGVCRESLAFETAKGPEPVSIQRAEALYLSVYAKDASAEV  
VKPDQITSSNDDQIQKPVSSERLEKDLAKAIVKICGKMIPEAPEIECLSIGAREL  
PPPTVSGSSCVVMRRPKRLYIGOTDDELEGRIRAHRAKEGLQSSFLYLMVQGSMA  
COLETLINQLHEQGYSLANLADGKHRRNFGTSSLSLSDTVSVSL"  
37576, .39078  
/note="gene\_id:K7M2.10"  
/codon\_start=1  
/evidence=not experimental  
/product="beta-1,3-glucanase-like protein"  
/protein\_id="BAB02933.1"  
/db\_xref="GI:11994695"

CDS

/translation="WAGERSKLTNTNHNHQLIILCYFLIIISQVSIASNTSNVGNWG  
INASHQLPPEKVVKMDNSFTKLFEADQNIIDLALIGSDIEVMIGIPNRFLEMAQ  
DTSVAASVEENVTAYSYNGVNIKIYIAGNEPFLQYNTGYFTVFLSHDINIQRALE  
EADLKNKVYVPFNADIVFSPEANVPVPSAGDPRELDRATIEIINFLYSLSPFTVNI  
YFPLSLYGNAYFLPDAFDFGTNKSRLDGNLVYTNVFDANLDTLICAMERYSLGMIK  
TVGEVGNVTDGDKNANVSAKRPFNGQMVKHAMSGNGTTPARKGVINDVYLFSLVDEDAK

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 11:40:28 ; Search time 742 Seconds  
(without alignments)  
11130.050 Million cell updates/sec

Title: US-10-086-464-1  
Perfect score: 1944  
Sequence: 1 atgtcctcgccgcgtctcc.....atagtggaccttctcttaa 1944

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1944	100.0	1944	4	AAF77094 Brassica
2	1944	100.0	1944	9	ACF36548 B. napus
3	1944	100.0	2189	9	ACF36549 B. napus
4	66	3.4	1959	9	ACF36557 A. thalia
5	47	2.4	2324	3	AAC36968 Arabidops
6	26	1.3	1515	9	ACF36555 Z. mays P
7	26	1.3	110000	3	Continuation (7 of
8	26	1.3	110000	3	Continuation (8 of
9	24	1.2	1389	7	ADA70853 Rice gene
10	24	1.2	1515	8	ADA27371 HPV-16 L1
11	24	1.2	1515	8	ADA92549 HPV-16 L1
12	24	1.2	1515	8	ADA14298 HPV-16 L1
13	24	1.2	1515	8	ADA58571 HPV-16 L1
14	23	1.2	897	7	AAD47597 Human HGF
15	23	1.2	897	7	AAD47600 Human HGF
16	23	1.2	897	7	AAD47598 Human HGF
17	23	1.2	1005	7	ACF70783 Phototrab
18	23	1.2	1281	7	AD47599 Human HGF
19	23	1.2	1572	5	AAS84463 DNA encod
20	23	1.2	3252	9	ADB61530 Hepatocyt
21	23	1.2	4173	9	ADB61585 Hepatocyt
22	23	1.2	4517	6	ABV99385 Human NOV
23	23	1.2	4586	7	AAD48130 Human C-M

24	23	1.2	4620	7	AAD47594	Human hep
25	23	1.2	4620	9	ADC78858	Human PRO
26	23	1.2	4626	2	AAV20444	Human c-m
27	23	1.2	110000	7	ACF67367_38	Continuation (39 o
28	23	1.2	110000	7	ACF65388_09	Continuation (10 o
29	21	1.1	1119	6	AAD30797	Hansenula
30	21	1.1	1269	2	AAD47060	ZmPK1 hom
31	21	1.1	1949	4	AAF44708	Novel pro
32	21	1.1	2048	6	ABN59741	Novel hum
33	21	1.1	2880	4	AAF77095	Arabidops
34	21	1.1	2880	9	ACF36550	Arabidops
35	21	1.1	2891	5	AAD07043	Human DNA
36	21	1.1	50959	3	AAF22304	Arabidops
37	21	1.1	72592	3	AAF22300	BAC conta
38	21	1.1	222880	9	ADC87622	Human GPC
39	21	1.1	349938	9	ADC87621	Human GPC
40	20	1.0	213	6	ABL86637	Human ova
41	20	1.0	432	7	ABX62007	Arabidops
42	20	1.0	609	4	AAF97928	Human sec
43	20	1.0	657	9	ADB95095	A. thalia
44	20	1.0	772	3	AAC49631	Arabidops
45	20	1.0	1002	6	AAD44411	Human hun

## ALIGNMENTS

## RESULT 1

ID	AAF77094	standard; DNA; 1944 BP.
AC	AAF77094;	
XX		
DT	17-MAY-2001	(first entry)
XX		
DE	Brassica napus PERK1 DNA.	
XX		
KW	Proline-rich extensin-like receptor kinase; PERK; resistance; plant; ds.	
XX		
OS	Brassica napus.	
XX		
FN	WO200114563-Al.	
XX		
PD	01-MAR-2001.	
XX		
PF	18-AUG-2000; 2000WO-CA000966.	
XX		
PR	19-AUG-1999; 99US-0149466P.	
PR	13-OCT-1999; 99US-0159122P.	
XX		
PA	(GORI/) GORING D.	
PA	(SILV/) SILVA N.	
XX		
PI	Goring D, Silva N;	
XX		
DR	WPI; 2001-244305/25.	
XX		
PT	New proline-rich, extensin-like receptor kinase nucleic acids and	
PT	polypeptides useful for increasing plant wounding or pathogen resistance,	
PT	or for producing transgenic plants with increased wounding or pathogen	
PT	resistance.	
XX		
PS	Claim 6; Fig 1; 91pp; English.	
XX		
CC	The present invention relates to proline-rich extensin-like receptor	
CC	kinase (PERK). The PERK nucleic acids and polypeptides are useful for	
CC	increasing the resistance of plants to wounding and pathogens. These are	
CC	also useful for producing transgenic plants with increased wounding and	
CC	pathogen resistance compared with a wild type plant, as well as in assays	
CC	for identifying and developing compounds to inhibit and/or enhance	
XX	polypeptide function directly	
SQ	Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;	



DE	B. napus PERK1 receptor kinase encoding cDNA.	
XX		
KW	PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;	
KW	wound; pathogen resistance; plant growth; seed production; gene; ss.	
XX		
OS	Brassica napus.	
XX		
-FH	Key	Location/Qualifiers
FT	CDS	1..1944
FT		/*tag= a
FT		/transl_except= (pos: 1332..1334, aa: Phe)
FT		/product= "PERK1"
FT		/note= "Proline-rich Extensin-like Receptor Kinase"
XX		
XX	WO2003072763-A1.	
PN		
XX		
PD	04-SEP-2003.	
XX		
XX	28-FEB-2003; 2003WO-CA000274.	
XX		
PR	28-FEB-2002; 2002CA-02373903.	
PR	28-FEB-2002; 2002US-00086464.	
XX		
PA	(GORI/) GORING D.	
PA	(SILV/) SILVA N.	
PA	(HAFF/) HAPFANI Y Z.	
XX		
PI	Goring D, Silva N, Haffani YZ;	
XX		
DR	WPI; 2003-712727/67.	
DR	P-PSDB; ABR82937.	
XX		
PT	Producing a transgenic plant having an increased plant resistance, plant growth or seed production comprises transforming a plant with a nucleic acid molecule having a Proline-rich Extensin-like Receptor Kinase activity.	
PT		
XX		
PS	Claim 3; Fig 1A; 123pp; English.	
XX		
CC	The invention relates to producing a transgenic plant having increased plant height, number of branches, number of seed pods and/or seed production compared to a non-transgenic plant, and/or quicker flowering or later senescence compared to a non-transgenic plant. The method involves transforming a plant with a vector including a Proline-rich Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic acid molecule having PERK activity. The method, as well as the PERK nucleic acid molecule and polypeptide, are useful in increasing plant resistance to wounding and pathogens and in increasing plant growth and seed production. The nucleic acid molecule and polypeptide may also be used in producing transgenic plants or transgenic host cells. The present sequence represents a cDNA encoding a B. napus PERK1 receptor kinase polypeptide	
XX		
SQ	Sequence 1944 BP; 464 A; 550 C; 459 G; 471 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 1944; DB 9; Length 1944;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ATGTCTCTCGGCGCGTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 60	
Db	1 ATGTCTCTCGGCGCGTCTCGGGGACTGGTTCGCTCCATCTCCACCATCAAACTCCACA 60	
Qy	61 ACCACCACTCTCTCCAGTTCGGCTCTCTCCGACACACACCTTCTCTCTCTCGCGCG 120	
Db	61 ACCACCACTCTCTCCAGTTCGGCTCTCTCCGACACACACCTTCTCTCTCTCGCGCG 120	
Qy	121 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180	
Db	121 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180	
Qy	181 CCATCTCCACCAACTCCATCTACGGCGGATCTCCACCTCTCTCTCTCTCTCTCTCCA 240	
Db	181 CCATCTCCACCAACTCCATCTACGGCGGATCTCCACCTCTCTCTCTCTCTCTCTCCA 240	

```
Qy 1321 TACTTGGCTCCGAAATACGCTGCAAGCGGAAGCTCACGGAGAAAGTCTGACGTTTTCTCA 1380
Db 1321 TACTTGGCTCCGAAATACGCTGCAAGCGGAAGCTCACGGAGAAAGTCTGACGTTTTCTCA 1380
Qy 1381 TTTGGCGTTGGCTTTGGAGCTCATTACTGGAGCTGCGACCGTTGATGCCAAATGTC 1440
Db 1381 TTTGGCGTTGGCTTTGGAGCTCATTACTGGAGCTGCGACCGTTGATGCCAAATGTC 1440
Qy 1441 TATGTAGATGACAGCTTAGTTGACTGGGCGACGACCATTTGCTTAACCGAGCATCTGAGCAA 1500
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGCGACGACCATTTGCTTAACCGAGCATCTGAGCAA 1500
Qy 1501 GGAGACTTTGAGGGTTAGCTGATGCAAAAGATGAATATGCGGTATGACAGAGAGGATG 1560
Db 1501 GGAGACTTTGAGGGTTAGCTGATGCAAAAGATGAATATGCGGTATGACAGAGAGGATG 1560
Qy 1561 GCTGCGATGTTGCTTGTGCTGCGGCTTGTGTCGCCATTTCAGCTCGCGCGAGACCTCGC 1620
Db 1561 GCTGCGATGTTGCTTGTGCTGCGGCTTGTGTCGCCATTTCAGCTCGCGCGAGACCTCGC 1620
Qy 1621 ATGAGCCAGATTGTGCGTGCCTTAGAAGAAATGTATCACTGTGATCTTTAAAGAAAGG 1680
Db 1621 ATGAGCCAGATTGTGCGTGCCTTAGAAGAAATGTATCACTGTGATCTTTAAAGAAAGG 1680
Qy 1681 ATGAGACAGCTCAAAGCAATGTATACAGCTCATACGAGGAAGCAGATTTAGACTCG 1740
Db 1681 ATGAGACAGCTCAAAGCAATGTATACAGCTCATACGAGGAAGCAGATTTAGACTCG 1740
Qy 1741 AGCCAGTACATGAGACATGAAGAGTTTATAGGAATGGCACTTGGAACTCAAGAGTAC 1800
Db 1741 AGCCAGTACATGAGACATGAAGAGTTTATAGGAATGGCACTTGGAACTCAAGAGTAC 1800
Qy 1801 AACCAGCGGTGAGTACAGTAATCCGACAGTGAATGACCTGATCCGCTCTGGTTCA 1860
Db 1801 AACCAGCGGTGAGTACAGTAATCCGACAGTGAATGACCTGATCCGCTCTGGTTCA 1860
Qy 1861 AGCAGGAGGCCAAACACACAGGAAATGAGATGGGGAAGATTAAGAGAACCGGTGAG 1920
Db 1861 AGCAGGAGGCCAAACACACAGGAAATGAGATGGGGAAGATTAAGAGAACCGGTGAG 1920
Qy 1921 GGTATAGTGACCTTCTCTTTAA 1944
Db 1921 GGTATAGTGACCTTCTCTTTAA 1944
```

```
RESULT 3
ACF36549
ID ACF36549 standard; cdna; 2189 BP.
XX
AC AC
XX AC
XX AC
DT 18-DEC-2003 (first entry)
XX
DE B. napus PERK1 receptor kinase encoding cdna.
XX
XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
XX wound; pathogen resistance; plant growth; seed production; gene; ss.
XX
OS Brassica napus.
XX
FH Key Location/Qualifiers
FT 1. .2189
FT CDS /*tag= c
FT /note= "ABR82937; this protein contains the amino acids
FT corresponding to 5' and 3'UTR regions, though only the
FT relevant aa residues (ABR82937) is used in the invention"
FT 1. .96
FT /*tag= a
FT 97. .2040
FT /*tag= b
FT /transl_except= (pos: 1332. .1334, aa: Phe)
FT /product= "PERK1"
FT 5'UTR
FT CDS
FT 1. .2189
FT CDS
FT 97. .2040
FT /*tag= b
FT /transl_except= (pos: 1332. .1334, aa: Phe)
FT /product= "PERK1"
FT 3'UTR
FT CDS
FT 1. .2189
FT CDS
FT 97. .2040
FT /*tag= d
FT /note= "ABR82937; proline-rich Extensin-like Receptor
FT Kinase"
FT 2041. .2189
FT /*tag= d
XX
XX W02003072763-A1.
XX
XX 04-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-CA000274.
XX
XX 28-FEB-2002; 2002CA-02373903.
XX 28-FEB-2002; 2002US-00086464.
XX
XX (GORI/) GORING D.
XX (SILV/) SILVA N.
XX (HAFF/) HAFFANI Y Z.
XX
XX Goring D, Silva N, Haffani YZ;
XX
XX WPI; 2003-712727/67.
XX P-PSDB; ABR82937, ABR82938.
XX
XX Producing a transgenic plant having an increased plant resistance, plant
XX growth or seed production comprises transforming a plant with a nucleic
XX acid molecule having a Proline-rich Extensin-like Receptor Kinase
XX activity.
XX
XX Disclosure; Fig 1D; 123pp; English.
XX
XX The invention relates to producing a transgenic plant having increased
XX plant height, number of branches, number of seed pods and/or seed
XX production compared to a non-transgenic plant, and/or quicker flowering
XX or later senescence compared to a non-transgenic plant. The method
XX involves transforming a plant with a vector including a Proline-rich
XX Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic
XX acid molecule having PERK activity. The method, as well as the PERK
XX nucleic acid molecule and polypeptide, are useful in increasing plant
XX resistance to wounding and pathogens and in increasing plant growth and
XX seed production. The nucleic acid molecule and polypeptide may also be
XX used in producing transgenic plants or transgenic host cells. The present
XX sequence represents a cdna encoding a B. napus PERK1 receptor kinase
XX polypeptide
XX
XX Sequence 2189 BP; 544 A; 592 C; 494 G; 559 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 1944; DB 9; Length 2189;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTCTCGCGCGGTCTCCGGGAGTGGTTCGCTCCATCTCCACCATCAAACTCCACA 60
Db 97 ATGTCTCGCGCGGTCTCCGGGAGTGGTTCGCTCCATCTCCACCATCAAACTCCACA 156
Qy 61 ACACAGACTCTCTCTCCAGCTTCGCTCCCTCCACCATCAAACTCCACA 120
Db 157 ACCACCACTCTCTCTCCAGCTTCGCTCCCTCCACCATCAAACTCCACA 216
Qy 121 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 217 CCATCCACTATTCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276
Qy 181 CCATCTCCACCACTCCATCTAGCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCTCCA 240
Db 277 CCATCTCCACCACTCCATCTAGCGCGGATCTCCACCTCTCTCTCTCTCTCTCTCCA 336
Qy 241 CCAGCTCAACTAGCGCGGATCTCCACCGGACCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 337 CCAGCTCAACTAGCGCGGATCTCCACCGGACCTCTCTCTCTCTCTCTCTCTCTCTCT 396
Qy 301 CCAGCTCAACTAGCGCGGATCTCCACCGGACCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 397 CCAGCTCAACTAGCGCGGATCTCCACCGGACCTCTCTCTCTCTCTCTCTCTCTCTCT 456
```

```
Qy 361 CCATCTTCTCCTCCGCGCGTCTCTCTTCCGACGGTTTATCAACAGGAGTGGTGGT 420
Db 457 CCATCTTCTCCTCCGCGCGTCTCTCTTCCGACGGTTTATCAACAGGAGTGGTGGT 516
Qy 421 GGAATCCCATCGGAGAGTGGTCTCTCTGTGTAGTAGTACTCTGATTTGTCTCTCTGT 480
Db 517 GGAATCCCATCGGAGAGTGGTCTCTCTGTGTAGTAGTACTCTGATTTGTCTCTCTGT 576
Qy 481 AAGNAGAAACGACGGAGAGAGAAAGATGCTTACTATGTTCTCCGCCACCTCTCTCT 540
Db 577 AAGNAGAAACGACGGAGAGAGAAAGATGCTTACTATGTTCTCCGCCACCTCTCTCT 636
Qy 541 GGTCCCAAGCCGAGAGACCTTACGGTGGACAGCAACAAATGGCGGCAACAAACGCA 600
Db 637 GGTCCCAAGCCGAGAGACCTTACGGTGGACAGCAACAAATGGCGGCAACAAACGCA 696
Qy 601 ACACCAACCGTCAGATCATGTGTCAGTCACTACCAACCAACCTAAAGGTCATCTCCA 660
Db 697 ACACCAACCGTCAGATCATGTGTCAGTCACTACCAACCAACCTAAAGGTCATCTCCA 756
Qy 661 CCAGGCAACCTCTCCACCTCCACACCGGCTTTCATGACGACGAGGGGCTCCGAC 720
Db 757 CCAGGCAACCTCTCCACCTCCACACCGGCTTTCATGACGACGAGGGGCTCCGAC 816
Qy 721 TACTCGGACCGTCAGTCTTCTCCACCGTCTCCAGGGCTTGTGTTAGGCTTCTCCAAA 780
Db 817 TACTCGGACCGTCAGTCTTCTCCACCGTCTCCAGGGCTTGTGTTAGGCTTCTCCAAA 876
Qy 781 AGCATTTTCATACGAGGAGCTAGCTAGAGCCACCAATGGTTTCTCCGAGGGCAACTTG 840
Db 877 AGCATTTTCATACGAGGAGCTAGCTAGAGCCACCAATGGTTTCTCCGAGGGCAACTTG 936
Qy 841 TTAGGACAAAGCGGTTTCGGTTAGCTGCAACAAAGGTGTGTCCTAGTGGGAAAGAGTT 900
Db 937 TTAGGACAAAGCGGTTTCGGTTAGCTGCAACAAAGGTGTGTCCTAGTGGGAAAGAGTT 996
Qy 901 GCTGTGAAGCAGTTGAAGTTGGAGTGGTCAGGGAGAGGGAGTTTCAGGCAAGAGTT 960
Db 997 GCTGTGAAGCAGTTGAAGTTGGAGTGGTCAGGGAGAGGGAGTTTCAGGCAAGAGTT 1056
Qy 961 GAGATCATGACGAGTTTACCACAGGCACTGGTGTCTCTTGTGTTATTTGCATCGCC 1020
Db 1057 GAGATCATGACGAGTTTACCACAGGCACTGGTGTCTCTTGTGTTATTTGCATCGCC 1116
Qy 1021 GGTGCGAAAAGATTGCTGTATGAGTTTGTCTTAACAAACAACTCTCGAGCTTCACCTC 1080
Db 1117 GGTGCGAAAAGATTGCTGTATGAGTTTGTCTTAACAAACAACTCTCGAGCTTCACCTC 1176
Qy 1081 CATGGCGAGGAGCGGCTCAATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1140
Db 1177 CATGGCGAGGAGCGGCTCAATGGAATGGAGCACCAGATTGAAGATTGCTCTTGGATCT 1236
Qy 1141 GCTAAAGGACTTTCTTATCTTTCATGAAGATTGCAATCTTAAATCATTCACCGTGATATC 1200
Db 1237 GCTAAAGGACTTTCTTATCTTTCATGAAGATTGCAATCTTAAATCATTCACCGTGATATC 1296
Qy 1201 AAGCTTTCAACATATTTGATAGATTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT 1260
Db 1297 AAGCTTTCAACATATTTGATAGATTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTT 1356
Qy 1261 GCTAAGATTGCTTCTGTATCAACACCGCATGTATCAACAGTGTGATGGAACTTTGGG 1320
Db 1357 GCTAAGATTGCTTCTGTATCAACACCGCATGTATCAACAGTGTGATGGAACTTTGGG 1416
Qy 1321 TACTTGGCTCCGGAATACGCTGCAAGCGGAAAGCTCACGGAGAGTCTGACGTTTCTCA 1380
Db 1417 TACTTGGCTCCGGAATACGCTGCAAGCGGAAAGCTCACGGAGAGTCTGACGTTTCTCA 1476
Qy 1381 TTTGGCTGTGCTTTTGGAGCTCATTTAGGAGTGCACCCGTTGATGCCAAATGTC 1440
Db 1477 TTTGGCTGTGCTTTTGGAGCTCATTTAGGAGTGCACCCGTTGATGCCAAATGTC 1536
```

```
Qy 1441 TATGTAGATGACAGCTTAGTGTGCTGGCCACGACCATTTGCTTAACCGAGCATCTGAGCAA 1500
Db 1537 TATGTAGATGACAGCTTAGTGTGCTGGCCACGACCATTTGCTTAACCGAGCATCTGAGCAA 1596
Qy 1501 GGAGACTTTGAGGGTTTAGCTGATGCAAAGATGAATAATGGGTATGAACAGAGGAGATG 1560
Db 1597 GGAGACTTTGAGGGTTTAGCTGATGCAAAGATGAATAATGGGTATGAACAGAGGAGATG 1656
Qy 1561 GCTCGCATGGTTTCTGCTGCTGGCTTGTGTCGCCATTCAGCTCCGCCAGACCTCGC 1620
Db 1657 GCTCGCATGGTTTCTGCTGCTGGCTTGTGTCGCCATTCAGCTCCGCCAGACCTCGC 1716
Qy 1621 ATGAGCCAGATTGCTGCTGCTTAGAAGGAAATGTATCACTGTGTCAGATCTTAACGAAAGG 1680
Db 1717 ATGAGCCAGATTGCTGCTGCTTAGAAGGAAATGTATCACTGTGTCAGATCTTAACGAAAGG 1776
Qy 1681 ATGAGCCAGGTCAAGCAATGTATACAGTCTATACGGAGAAAGCAACCGATTATGACTCG 1740
Db 1777 ATGAGCCAGGTCAAGCAATGTATACAGTCTATACGGAGAAAGCAACCGATTATGACTCG 1836
Qy 1741 AGCCAGTACAAATGAAGACATGAAGAGTTTGTAGGAAATGGCACTTGAACCTCAAGAGTAC 1800
Db 1837 AGCCAGTACAAATGAAGACATGAAGAGTTTGTAGGAAATGGCACTTGAACCTCAAGAGTAC 1896
Qy 1801 AACGCCACGGGTGAGTACAGTAATCCGACCACTGACTATGGAAGTGTACCCGCTCTGGTTCA 1860
Db 1897 AACGCCACGGGTGAGTACAGTAATCCGACCACTGACTATGGAAGTGTACCCGCTCTGGTTCA 1956
Qy 1861 AGCAGGAGGGCCAAACCAACACCGCAAAATGGAGATGGGGAAGATTAAAGAGAACCGGTGAC 1920
Db 1957 AGCAGGAGGGCCAAACCAACACCGCAAAATGGAGATGGGGAAGATTAAAGAGAACCGGTGAC 2016
Qy 1921 GGTATAGTGGACCTTCTCTTTAA 1944
Db 2017 GGTATAGTGGACCTTCTCTTTAA 2040
```

## RESULT 4

ACF36557  
ID ACF36557 standard; DNA; 1959 BP.

XX ACF36557;

XX AC

XX 18-DEC-2003 (first entry)

XX DE A. thaliana PERK1 protein encoding genomic DNA.

XX XX PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;  
XX KM wound; pathogen resistance; plant growth; seed production; gene; ds.

XX XX Arabidopsis thaliana.

XX OS

XX WO2003072763-A1.

XX PN

XX 04-SEP-2003.

XX XX

XX 28-FEB-2003; 2003WO-CA000274.

XX XX

XX 28-FEB-2002; 2002CA-02373903.

XX XX

XX 28-FEB-2002; 2002US-00086464.

XX XX (GORI/) GORING D.

XX PA (SILV/) SILVA N.

XX PA (HAFF/) HAFFANI Y Z.

XX XX Goring D, Silva N, Haffani YZ;

XX XX WPI; 2003-712727/67.

XX XX Producing a transgenic plant having an increased plant resistance, plant  
PT growth or seed production comprises transforming a plant with a nucleic  
PT acid molecule having a Proline-rich Extensin-like Receptor Kinase  
PT activity.



PR 22-JUL-1999; 99US-01451192P.  
 PR 23-JUL-1999; 99US-01451145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145226P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 28-JUL-1999; 99US-0145918P.  
 PR 02-AUG-1999; 99US-0146336P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148117P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148694P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 23-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 28-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 18-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.

PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161933P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 2.4%; Score 47; DB 3; Length 2324;

Best Local Similarity 100.0%; Pred. No. 1.7e-11; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1585 GCTTGTTCGCCATTGAGCTCGCCGAGACCTCGCATGAGCCAGAT 1631  
 |||||  
 Db 1700 GCTTGTTCGCCATTGAGCTCGCCGAGACCTCGCATGAGCCAGAT 1746

RESULT 6

ACF36555  
 ID ACF36555 standard; DNA: 1515 BP.

XX ACF36555;

DT 18-DEC-2003 (first entry)

DE Z. mays PERK protein encoding genomic DNA.

KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;  
 KW wound; pathogen resistance; plant growth; seed production; maize; gene;  
 ds.

XX Zea mays.

XX WO2003072763-A1.

XX 04-SEP-2003.

XX 28-FEB-2003; 2003WO-CA000274.

XX 28-FEB-2002; 2002CA-02373903.

XX 28-FEB-2002; 2002US-00086464.

XX (GORI/) GORING D.

XX (SILV/) SILVA N.

XX (HAFF/) HAFFANI Y Z.

XX Goring D, Silva N, Haffani YZ;

XX WPI; 2003-712727/67.

PT Producing a transgenic plant having an increased plant resistance, plant growth or seed production comprises transforming a plant with a nucleic acid molecule having a Proline-rich Extensin-like Receptor Kinase activity.

PS Disclosure; Fig 21b; 123pp; English.

CC The invention relates to producing a transgenic plant having increased plant height, number of branches, number of seed pods and/or seed production compared to a non-transgenic plant, and/or quicker flowering or later senescence compared to a non-transgenic plant. The method involves transforming a plant with a vector including a proline-rich Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic acid molecule having PERK activity. The method, as well as the PERK nucleic acid molecule and polypeptide, are useful in increasing plant resistance to wounding and pathogens and in increasing plant growth and

CC seed production. The nucleic acid molecule and polypeptide may also be  
CC used in producing transgenic plants or transgenic host cells. The present  
CC sequence represents a Z. mays PERK protein encoding genomic DNA (TIGR  
CC Accession No. AY108241)

XX Sequence 1515 BP; 400 A; 330 C; 381 G; 404 T; 0 U; 0 Other;

Query Match 1.3%; Score 26; DB 9; Length 1515;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1231 TTTGAAGCTAAGTGTGCTGATTGG 1256  
Db 601 TTTGAAGCTAAGTGTGCTGATTGG 626

## RESULT 7

AAF22305\_06/c  
Continuation (7 of 11) of AAF22305 from base 600001 (Arabidopsis thaliana chromosome 4 c  
WP Sequence split into 11 fragments LOCUS AAF22305 Accession Aaf22305  
WP Fragment Name Begin End  
WP AAF22305\_00 1 110000  
WP AAF22305\_01 100001 210000  
WP AAF22305\_02 200001 310000  
WP AAF22305\_03 300001 410000  
WP AAF22305\_04 400001 510000  
WP AAF22305\_05 500001 610000  
WP AAF22305\_06 600001 710000  
WP AAF22305\_07 700001 810000  
WP AAF22305\_08 800001 910000  
WP AAF22305\_09 900001 1010000  
WP AAF22305\_10 1000001 1082138

Query Match 1.3%; Score 26; DB 3; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 CTTGCTATGAGTTGTTCTTAACAA 1061  
Db 109643 CTTGCTATGAGTTGTTCTTAACAA 109618

## RESULT 8

AAF22305\_07/c  
Continuation (8 of 11) of AAF22305 from base 700001 (Arabidopsis thaliana chromosome 4 c  
WP Sequence split into 11 fragments LOCUS AAF22305 Accession Aaf22305  
WP Fragment Name Begin End  
WP AAF22305\_00 1 110000  
WP AAF22305\_01 100001 210000  
WP AAF22305\_02 200001 310000  
WP AAF22305\_03 300001 410000  
WP AAF22305\_04 400001 510000  
WP AAF22305\_05 500001 610000  
WP AAF22305\_06 600001 710000  
WP AAF22305\_07 700001 810000  
WP AAF22305\_08 800001 910000  
WP AAF22305\_09 900001 1010000  
WP AAF22305\_10 1000001 1082138

Query Match 1.3%; Score 26; DB 3; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 CTTGCTATGAGTTGTTCTTAACAA 1061  
Db 9643 CTTGCTATGAGTTGTTCTTAACAA 9618

## RESULT 9

ADA70853/c  
ID ADA70853 standard; DNA; 1389 BP.  
XX  
AC ADA70853;

XX 20-NOV-2003 (first entry)  
DT Rice gene, SEQ ID 4176.  
DE

XX Plant; bacterial infection; fungal infection; viral infection; rice;  
KW gene; ds.  
KW

OS Oryza sativa.

XX WC2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.

PS Claim 6; SEQ ID NO 4176; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

XX Sequence 1389 BP; 412 A; 278 C; 406 G; 293 T; 0 U; 0 Other;

Query Match 1.2%; Score 24; DB 7; Length 1389;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 670 CCTCTCCAGCTCCACCACCGCCT 693  
Db 115 CCTCTCCAGCTCCACCACCGCCT 92

## RESULT 10

ADA27371  
ID ADA27371 standard; DNA; 1515 BP.

XX ADA27371;

XX 20-NOV-2003 (first entry)

XX HPV-16 L1 codon optimised DNA sequence SEQ ID NO:1.

XX cell line; American Type Culture Collection PTA-4047; ATCC-4047;  
KW baculoviruses; viral recombinant protein; virus-like particle; vaccine;  
KW diagnostic reagent; human papillomavirus type 16; HPV-16; L1; gene; ds.

XX Synthetic.  
OS Human papillomavirus type 16.  
XX Key Location/Qualifiers  
XX CDS 1. .1515

```
FT      /*tag= a
FT      /product= "HPV-16 L1 protein"
FT      /transl_except= (pos:208..210,aa:Tyr)
FT      /transl_except= (pos:295..297,aa:Val)
XX
XX      WO2003068804-A2.
XX
XX      21-AUG-2003.
XX
XX      14-FEB-2003; 2003WO-US004516.
XX
XX      14-FEB-2002; 2002US-0356113P.
XX      14-FEB-2002; 2002US-0356118P.
XX      14-FEB-2002; 2002US-0356119P.
XX      14-FEB-2002; 2002US-0356123P.
XX      14-FEB-2002; 2002US-0356126P.
XX      14-FEB-2002; 2002US-0356133P.
XX      14-FEB-2002; 2002US-0356135P.
XX      14-FEB-2002; 2002US-0356136P.
XX      14-FEB-2002; 2002US-0356151P.
XX      14-FEB-2002; 2002US-0356152P.
XX      14-FEB-2002; 2002US-0356154P.
XX      14-FEB-2002; 2002US-0356156P.
XX      14-FEB-2002; 2002US-0356157P.
XX      14-FEB-2002; 2002US-0356161P.
XX      14-FEB-2002; 2002US-0356162P.
XX
XX      (NOVA-) NOVAVAX INC.
XX
XX      Robinson RA;
XX
XX      WPI; 2003-646475/61.
XX      P-PSDB; ADA27363.
XX
XX      New insect cell line designated ATCC PTA-4047, useful for replicating
XX      baculoviruses to produce large amounts of recombinant proteins of
XX      medical, pharmaceutical and veterinary importance.
XX
XX      Claim 10; Fig 1; 63pp; English.
XX
XX      The present invention describes a cell line comprising a cell that is a
XX      clone, derivative, mutant and/or transfectant of a cell line designated
XX      American Type Culture Collection (ATCC) PTA-4047. The cell upon culture
XX      grows continuously and retains the identifying characteristics of the
XX      cell line designated ATCC-4047. Also described is a process of making a
XX      cell line. The insect cell line is useful in replicating baculoviruses,
XX      as a host substrate for baculovirus plaque assays, as a source of insect
XX      proteins, acts as a depot for cell transfection to produce recombinant
XX      baculoviruses, and in expressing viral recombinant proteins.
XX      Extracellular and intracellular viral recombinant proteins and virus-like
XX      particles expressed from the cell line are useful as pharmaceutical
XX      compositions, vaccines or diagnostic reagents. The present sequence
XX      represents a human papillomavirus type 16 (HPV-16) L1 codon optimised
XX      sequence, which is used in the exemplification of the present invention.
XX
XX      Sequence 1515 BP; 297 A; 532 C; 308 G; 278 T; 0 U; 0 Other;
XX
XX      Query Match      1.2%; Score 24; DB 8; Length 1515;
XX      Best Local Similarity 100.0%; Pred. No. 1;
XX      Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      669 ACCTCTCCACCTCCACCGGCC 692
XX      |||||
XX      Db      1471 ACCTCTCCACCTCCACCGGCC 1494
XX
XX      RESULT 11
XX      ADA92549
XX      ID      ADA92549 standard; DNA; 1515 BP.
XX
XX      AC      ADA92549;
XX
XX      XX
XX      DT      20-NOV-2003 (first entry)
```

```
XX      HPV-16 L1 codon optimised DNA sequence SEQ ID NO:1.
XX
XX      codon optimised; viral capsid protein; virus-like particle; VLP;
XX      antigenic; human papillomavirus infection; virucide; vaccine;
XX      gene therapy; human papillomavirus type 16; dysplasia; infection; HPV-16;
XX      L1; gene; ds.
XX
XX      Synthetic.
XX      OS      Human papillomavirus type 16.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..1515
XX              /*tag= a
XX              /product= "codon optimised L1 amino acid sequence"
XX              /transl_except= (pos:208..210,aa:Tyr)
XX              /transl_except= (pos:295..297,aa:Val)
XX
XX      WO2003068933-A2.
XX
XX      21-AUG-2003.
XX
XX      14-FEB-2003; 2003WO-US004480.
XX
XX      14-FEB-2002; 2002US-0356113P.
XX      14-FEB-2002; 2002US-0356118P.
XX      14-FEB-2002; 2002US-0356119P.
XX      14-FEB-2002; 2002US-0356123P.
XX      14-FEB-2002; 2002US-0356126P.
XX      14-FEB-2002; 2002US-0356133P.
XX      14-FEB-2002; 2002US-0356135P.
XX      14-FEB-2002; 2002US-0356150P.
XX      14-FEB-2002; 2002US-0356151P.
XX      14-FEB-2002; 2002US-0356152P.
XX      14-FEB-2002; 2002US-0356154P.
XX      14-FEB-2002; 2002US-0356156P.
XX      14-FEB-2002; 2002US-0356157P.
XX      14-FEB-2002; 2002US-0356161P.
XX      14-FEB-2002; 2002US-0356162P.
XX
XX      (NOVA-) NOVAVAX INC.
XX
XX      Robinson RA;
XX
XX      WPI; 2003-689664/65.
XX      P-PSDB; ADA92541.
XX
XX      New codon optimized polynucleotide encoding a viral capsid protein that
XX      self assembles into a virus-like particle, useful for diagnosing,
XX      preventing or treating human papillomavirus infections or associated
XX      disorders.
XX
XX      Claim 7; Page 113; 123pp; English.
XX
XX      The present invention describes a codon optimised polynucleotide encoding
XX      a viral capsid protein that self assembles into a virus-like particle
XX      (VLP) that exhibits conformational antigenic epitopes capable of raising
XX      neutralising antibodies, where the VLP is expressed from a host cell
XX      extracellularly. Also described: (1) a vector comprising the above codon
XX      optimised polynucleotide operably linked to a eukaryotic or prokaryotic
XX      regulatory control element, capable of replication in prokaryotic and/or
XX      eukaryotic host; (2) a host cell comprising the vector; (3) a
XX      pharmaceutical or vaccine composition for treating, ameliorating or
XX      preventing a papillomavirus related disease or disorder, comprising a
XX      multiplicity of VLPs that exhibit conformational antigenic epitopes, and
XX      a carrier, diluent or adjuvant; (4) a diagnostic kit for detecting a
XX      papillomavirus infection, comprising a multiplicity of VLPs that exhibit
XX      conformational antigenic epitopes, and a detection agent comprising a
XX      detectable label; (5) a method for preparing the above codon optimised
XX      polynucleotide, comprising replacing codons that are underutilised in
XX      insect cells with codons that are utilised at high levels in insect
XX      cells, to create an initially-modified nucleotide sequence, and modifying
XX      the initially-modified nucleotide sequence by choosing a preferred codon
```

CC for the initially-modified sequence, where the ratio of GC nucleotide  
 CC pairs to AT nucleotide pairs in the further-modified nucleotide sequence  
 CC trends towards about 1:1, where the number of palindromic and stem-loop  
 CC DNA structures in the further-modified nucleotide sequence is minimised,  
 CC and where the number of transcription and post-transcription repressor  
 CC elements are minimised; and (6) methods for treating, ameliorating or  
 CC preventing a papillomavirus related disease or disorder, or for  
 CC protecting an individual against a papillomavirus infection, comprising  
 CC administering to an individual an amount of the composition or vaccine  
 CC cited above. The VLP has virucide activity and can be used in vaccines  
 CC and in gene therapy. The composition and methods of the present invention  
 CC are useful in diagnosing, preventing or treating human papillomavirus  
 CC infections or associated disorders, such as dysplasia. The present  
 CC sequence represents an HPV-16 codon optimised L1 nucleotide sequence from  
 CC the present invention.

SQ Sequence 1515 BP; 297 A; 632 C; 308 G; 278 T; 0 U; 0 Other;  
 Query Match 1.2%; Score 24; DB 8; Length 1515;  
 Best Local Similarity 100.0%; Pred. No. 1;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ACCTCTCCACCTCCACACCGCC 692  
 Db 1471 ACCTCTCCACCTCCACACCGCC 1494

## RESULT 12

AD14298  
 ID ADA14298 standard; DNA; 1515 BP.

XX ADA14298;

DT 06-NOV-2003 (first entry)

DE HPV-16 L1 codon optimised sequence SEQ ID NO:1.

KW purification; recombinant extracellular virus-like particle;  
 KW recombinant intracellular virus-like particle; virus-like particle; VLP;  
 KW virucide; vaccine; gene therapy; human papillomavirus; HPV; infection;  
 KW dysplasia; HPV-16; L1; codon optimised; gene; ds.

OS Human papillomavirus type 16.

XX Key Location/Qualifiers  
 FH 1. .1515  
 FT /\*tag= a  
 FT /product= "HPV-16 L1 protein"  
 FT /transl\_except= (pos:208. .210,aa:Tyr)  
 FT /transl\_except= (pos:295. .297,aa:Val)

XX WO2003068993-A1.

XX 21-AUG-2003.

XX 14-FEB-2003; 2003WO-US004474.

XX 14-FEB-2002; 2002US-0356113P.

XX 14-FEB-2002; 2002US-0356118P.

XX 14-FEB-2002; 2002US-0356119P.

XX 14-FEB-2002; 2002US-0356123P.

XX 14-FEB-2002; 2002US-0356126P.

XX 14-FEB-2002; 2002US-0356133P.

XX 14-FEB-2002; 2002US-0356135P.

XX 14-FEB-2002; 2002US-0356150P.

XX 14-FEB-2002; 2002US-0356151P.

XX 14-FEB-2002; 2002US-0356152P.

XX 14-FEB-2002; 2002US-0356154P.

XX 14-FEB-2002; 2002US-0356156P.

XX 14-FEB-2002; 2002US-0356157P.

XX 14-FEB-2002; 2002US-0356161P.

XX 14-FEB-2002; 2002US-0356162P.

PA (NOVA-) NOVAVAX INC.

XX Robinson RA, Thompson MW;

PI WPI; 2003-679645/64.

DR P-PSDB; ADA14290.

XX

PT Purifying a recombinant human papillomavirus (HPV) L1, useful for  
 PT diagnosing, preventing or treating HPV infections, comprise clarifying,  
 PT concentrating and diafiltering cells containing HPV particles.

PS Example 4; Fig 1A-C; 111pp; English.

XX The present invention describes a method for purifying a recombinant  
 CC extracellular or intracellular virus-like particle (VLP). The method  
 CC comprises harvesting a cell suspension comprising cells containing a  
 CC plurality of VLPs to produce a harvested supernatant, optionally  
 CC disrupting the harvested cells to produce cell lysates containing the  
 CC VLP, clarifying the harvested supernatant, concentrating the clarified  
 CC supernatant, diafiltering the concentrated supernatant, and recovering  
 CC the purified recombinant VLP. Also described: (1) a cell line designated  
 CC as SF-9S deposited as American Type Culture Collection (ATCC) PTA-4047;  
 CC (2) producing the cell line described above; (3) host cells that express  
 CC one or more recombinant gene products with an enhanced yield; (4)  
 CC producing a foreign protein in an insect cell; (5) nucleic acid sequences  
 CC that correspond to and code for human papillomavirus (HPV) polypeptides;  
 CC and (6) pharmaceutical compositions comprising an amount of the  
 CC recombinant viral gene products, VLPs, agonists, antagonists, or the  
 CC active fragment of a viral gene product. The VLPs have virucide activity,  
 CC and can be used in vaccines and in gene therapy. The method is useful in  
 CC isolating and purifying expressed viral gene products, including VLPs, in  
 CC vitro. The gene products or particles may be used in detecting,  
 CC preventing or treating HPV infections and associated symptoms like  
 CC dysplasia. The present sequence represents an HPV-16 L1 codon optimised  
 CC nucleotide sequence, which is given in the exemplification of the present  
 CC invention.

SQ Sequence 1515 BP; 297 A; 632 C; 308 G; 278 T; 0 U; 0 Other;

Query Match 1.2%; Score 24; DB 8; Length 1515;

Best Local Similarity 100.0%; Pred. No. 1;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 669 ACCTCTCCACCTCCACACCGCC 692

Db 1471 ACCTCTCCACCTCCACACCGCC 1494

## RESULT 13

AA58571

ID AA58571 standard; DNA; 1515 BP.

XX

XX AA58571;

XX

DT 04-DEC-2003 (first entry)

XX

DE HPV-16 L1 codon optimised DNA.

XX

Vaccine; humoral immunity; cell-mediated immunity; gene therapy; HPV;  
 virucide; papillomavirus infection; gene; Human papillomavirus; ds.

XX

OS Human papillomavirus.

XX

FH Key Location/Qualifiers

CDS 1. .1515

FT /\*tag= a

FT /product= "HPV-16 L1 protein"

FT /transl\_except= (pos:208. .210, aa:Tyr)

FT /transl\_except= (pos:295. .297, aa:Val)

FT /note= "No stop codon"

XX /partial

XX WO2003068163-A2.

```
XX PD 21-AUG-2003.
XX PF 14-FEB-2003; 2003WO-US004473.
XX PR 14-FEB-2002; 2002US-0356113P.
XX PR 14-FEB-2002; 2002US-0356118P.
XX PR 14-FEB-2002; 2002US-0356119P.
XX PR 14-FEB-2002; 2002US-0356123P.
XX PR 14-FEB-2002; 2002US-0356126P.
XX PR 14-FEB-2002; 2002US-0356133P.
XX PR 14-FEB-2002; 2002US-0356135P.
XX PR 14-FEB-2002; 2002US-0356150P.
XX PR 14-FEB-2002; 2002US-0356151P.
XX PR 14-FEB-2002; 2002US-0356152P.
XX PR 14-FEB-2002; 2002US-0356154P.
XX PR 14-FEB-2002; 2002US-0356156P.
XX PR 14-FEB-2002; 2002US-0356157P.
XX PR 14-FEB-2002; 2002US-0356161P.
XX PR 14-FEB-2002; 2002US-0356162P.
XX PA (NOVA-) NOVAVAX INC.
XX PI Robinson RA, Cioce V;
XX PF WPI; 2003-689598/65.
XX DR P-PSDB; AAE38615.
XX PS Claim 10; Fig 1A-1C; 126pp; English.
XX CC The present invention relates to chimeric virus-like particle comprising
CC a recombinant viral capsid protein that encapsulates a recombinant viral
CC protein during self assembly into a chimeric virus-like particle and
CC exhibiting conformational antigenic epitopes capable of eliciting
CC neutralising antibodies. The vaccine comprising the chimeric virus-like
CC particles are useful for inducing immunity (humoral and/or cell-mediated
CC immunity) against papillomavirus infection. The invention is also useful
CC in gene therapy. The present sequence is HPV (human papillomavirus)-16 L1
CC codon optimised DNA
XX SQ Sequence 1515 BP; 297 A; 632 C; 308 G; 278 T; 0 U; 0 Other;
Query Match 1.2%; Score 24; DB 8; Length 1515;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 669 ACCTCTCCACCTCCACCCGCC 692
DB 1471 ACCTCTCCACCTCCACCCGCC 1494
RESULT 14
AAD47597
ID AAD47597 standard; DNA; 897 BP.
XX AC
XX AAD47597;
XX DT
XX DE
XX 24-FEB-2003 (first entry)
XX DE Human HGFR M1250T mutant DNA fragment.
XX KW Human; hepatocyte growth factor receptor kinase; hHGRF; proto-oncogene;
XX protein co-ordinate data; drug design; hepatocyte growth factor; tumour;
XX therapy; enzyme; mutant; gene; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
1. .897
/*tag= a
/product= "Human HGFR mutant protein fragment"
```

```
FT CDS
1. .897
/*tag= a
/product= "Human HGFR mutant protein fragment"
/translation= (pos:424..426, aa:Ala)
/notes= "CDS does not include stop codon"
/partial
FT PN EP1243596-A2.
XX PD 25-SEP-2002.
XX PF 25-MAR-2002; 2002EP-00006616.
XX PR 23-MAR-2001; 2001US-0277968P.
XX PA (AGOU-) AGOURON PHARM INC.
XX PI Mroczkowski B, Hickey M, Mctigue MA, Murray BW, Parge H, Sarup J;
XX PI Zhu J;
XX DR WPI; 2003-001650/01.
XX DR P-PSDB; AAE30037.
XX CC New polynucleotide encoding human hepatocyte growth factor receptor,
XX useful for designing modulators of the receptor, and potential antitumor
XX agents.
XX PS Claim 4; Page 157; 204pp; English.
XX CC The invention relates to human hepatocyte growth factor receptor kinase
XX (hHGRF) and its nucleic acid sequence. The invention also relates to
XX methods for identification of hHGRF inhibitors. hHGRF DNA is used to
XX express the corresponding polypeptides and these are used to provide
XX protein co-ordinate data useful in rational drug design of compounds that
XX modulate kinase domain, potentially useful for treating diseases
XX associated with signalling between hHGRF and hepatocyte growth factor,
XX particularly tumours (since HGFR is a proto-oncogene). It can also be
XX used to search databases for related sequences and its fragments as
XX primers or probes. The present sequence is human HGFR mutant DNA fragment
XX used to illustrate the method of the invention
XX SQ Sequence 897 BP; 257 A; 193 C; 207 G; 240 T; 0 U; 0 Other;
Query Match 1.2%; Score 23; DB 7; Length 897;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1240 AAGGTTGCTGATTTGCTTTC 1262
DB 508 AAGGTTGCTGATTTGCTTTC 530
RESULT 15
AAD47600
ID AAD47600 standard; DNA; 897 BP.
XX AC
XX AAD47600;
XX DT
XX 24-FEB-2003 (first entry)
XX DE Human HGFR H1094R mutant DNA fragment.
XX KW Human; hepatocyte growth factor receptor kinase; hHGRF; proto-oncogene;
XX protein co-ordinate data; drug design; hepatocyte growth factor; tumour;
XX therapy; enzyme; mutant; gene; ds.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
1. .897
/*tag= a
/product= "Human HGFR mutant protein fragment"
```

FT /transl\_except= (pos:424, .426, aa:Ala)  
FT /transl\_except= (pos:667, .669, aa:Leu)  
FT /note= "CDS does not include stop codon"  
FT /partial

XX EF1243596-A2.

XX PD 25-SEP-2002.

XX XX 25-MAR-2002; 2002EP-00006516.

XX XX 23-MAR-2001; 2001US-0277968P.

XX PA (AGOU-) AGOURON PHARM INC.

XX Mroczkowski B, Hickey M, Metigue MA, Murray BW, Parge H, Sarup J;  
XX Zhu J;

XX DR WPI; 2003-001650/01.

XX DR P-PSDB; AAE30100.

XX New polynucleotide encoding human hepatocyte growth factor receptor,  
PT useful for designing modulators of the receptor, and potential antitumor  
XX agents.

XX PS Claim 6; Page 160-161; 204pp; English.

XX The invention relates to human hepatocyte growth factor receptor kinase  
CC (hGFR) and its nucleic acid sequence. The invention also relates to  
CC methods for identification of hGFR inhibitors. hGFR DNA is used to  
CC express the corresponding polypeptides and these are used to provide  
CC protein co-ordinate data useful in rational drug design of compounds that  
CC modulate kinase domain, potentially useful for treating diseases  
CC associated with signalling between hGFR and hepatocyte growth factor,  
CC particularly tumours (since HGFR is a proto-oncogene). It can also be  
CC used to search databases for related sequences and its fragments as  
CC primers or probes. The present sequence is human HGFR mutant DNA fragment  
XX used to illustrate the method of the invention

XX SQ Sequence 897 BP; 257 A; 191 C; 209 G; 240 T; 0 U; 0 Other;

Query Match 1.2%; Score 23; DB 7; Length 897;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1240 AAGTTGCTGATTTTGGTCTTGC 1262

Db 508 AAGTTGCTGATTTTGGTCTTGC 530

Search completed: May 12, 2004, 17:27:15  
Job time : 750 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:51:16 ; Search time 4545 Seconds  
(without alignments)  
12772.735 Million cell updates/sec

Title: US-10-086-464-1

Perfect score: 1944

Sequence: 1 atgcctcgccgcgtctcc.....atagtgacctctctttaa 1944

Scoring table: OLIGO\_NUC

Gapop\_60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	429	22.1	723	14	CD839231
2	273	14.0	453	28	BH442220 BOGLN68TR
3	187	9.6	753	28	BZ455058 BONKA33TF
4	65	3.3	314	14	CD811645

5	65	3.3	314	14	CD811912	CD811912 BN10.019G
c 6	53	2.7	561	9	AV539333	AV539333 AV539333
c 7	53	2.7	584	9	AV543493	AV543493 AV543493
8	42	2.2	754	28	BH432371	BH432371 BOGVQ93TR
9	32	1.6	342	28	BZ456327	BZ456327 BONEB66TR
c 10	30	1.5	744	28	BZ066106	BZ066106 1j170812.
11	29	1.5	505	12	BM061076	BM061076 KS01026B1
12	29	1.5	573	9	AV551753	AV551753 AV551753
13	29	1.5	607	29	CC968148	CC968148 BOJEB10TF
14	29	1.5	620	28	BH577077	BH577077 BOHCX77TR
15	29	1.5	651	13	BU043317	BU043317 PP_LE8001
16	29	1.5	695	28	BH928300	BH928300 Odf82F05.
c 17	28	1.4	713	12	BP175991	BP175991 BP175991
c 18	27	1.4	414	28	BH434903	BH434903 BOGVQ94TF
c 19	27	1.4	536	9	AL369682	AL369682 MCB32F07
c 20	26	1.3	204	14	CD004164	CD004164 VVA031D03
c 21	26	1.3	295	14	F14384	F14384 ATTS5296 AC
c 22	26	1.3	316	9	AV547061	AV547061 AV547061
c 23	26	1.3	360	9	AV534188	AV534188 AV534188
c 24	26	1.3	395	9	AV819894	AV819894 AV819894
c 25	26	1.3	399	9	AV797870	AV797870 AV797870
c 26	26	1.3	402	9	AV800600	AV800600 AV800600
c 27	26	1.3	411	9	AV799812	AV799812 AV799812
c 28	26	1.3	417	9	AV810018	AV810018 AV810018
c 29	26	1.3	421	9	AV788292	AV788292 AV788292
c 30	26	1.3	430	9	AV795981	AV795981 AV795981
c 31	26	1.3	436	9	AV793009	AV793009 AV793009
c 32	26	1.3	439	9	AV819129	AV819129 AV819129
c 33	26	1.3	459	9	AV538962	AV538962 AV538962
c 34	26	1.3	482	13	BU820923	BU820923 UB16CPE10
35	26	1.3	484	12	BG817458	BG817458 EM1_76 CO
36	26	1.3	593	10	BF176907	BF176907 EM1_4_B10
37	26	1.3	595	14	CF015663	CF015663 OBLAb11.x
38	26	1.3	613	14	CF232052	CF232052 PRAJX0000
39	26	1.3	637	13	BQ134241	BQ134241 1091014D0
40	26	1.3	641	14	CA248316	CA248316 SCCCFU509
41	26	1.3	645	13	BQ506869	BQ506869 EST614284
42	26	1.3	666	13	BU499754	BU499754 946178G10
43	26	1.3	673	14	CA164704	CA164704 SCSBR2312
44	26	1.3	682	13	CA095337	CA095337 SCCCL500
45	26	1.3	750	13	CA148443	CA148443 SCUFR2100

ALIGNMENTS

RESULT 1:  
CD839231  
LOCUS RFO2.114I04F010529 RFO2 Brassica napus cDNA clone RFO2114I04, mRNA  
DEFINITION sequence.  
ACCESSION CD839231  
VERSION CD839231.1 GI:32521171  
KEYWORDS EST.  
SOURCE Brassica napus (rape)  
ORGANISM Brassica napus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 723)  
AUTHORS Genoplate.  
TITLE Genoplate, a major partnership french program in plant genomics  
JOURNAL Unpublished (2003)  
COMMENT Contact: Genoplate  
Genoplate  
93, rue Henri Rochefort 91025 EVRY CEDEX France  
Tel: 33 1 69 47 54 00  
Fax: 33 1 69 47 54 10  
This sequence has been generated in the framework of the french  
plant genomics programme 'Genoplate' (<http://www.genoplate.com>)  
and <http://genoplate-info.infobiogen.fr>.

FEATURES  
source 1. .723

```
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="samourai (restored line)"
/db_xref="taxon:3708"
/clone="RF02114104"
/tissue_type="anthers"
/clone_lib="RF02"

ORIGIN
Query Match      22.1%; Score 429; DB 14; Length 723;
Best Local Similarity 99.6%; Pred. No. 2,7e-183;
Matches 529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1414 CGTGACCCGTTGATGCCAACAAATGTCTATGTAGATGACAGCTTAGTGTGCTGGCAGCA 1473
Db 71 CGTGACCCGTTGATGCCAACAAATGTCTATGTAGATGACAGCTTAGTGTGCTGGCAGCA 130

Qy 1474 CCATTGCTTAACCGAGCATCTGAGCAAGGACCTTTGAGGGTTTGTAGCTGATGCAAGATG 1533
Db 131 CCATTGCTTAACCGAGCATCTGAGCAAGGACCTTTGAGGGTTTGTAGCTGATGCAAGATG 190

Qy 1534 AATAATGGGTATGACAGAGAGAGATGCTCGCATGTTGCTTGTGCTGCGGCTTGTGTT 1593
Db 191 AATAATGGGTATGACAGAGAGAGATGCTCGCATGTTGCTTGTGCTGCGGCTTGTGTT 250

Qy 1594 CGCCATTGAGTGGCGGAGACCTCGCATGAGCCAGATTTGTCGCTTGTAGAGGAAAT 1653
Db 251 CGCCATTGAGTGGCGGAGACCTCGCATGAGCCAGATTTGTCGCTTGTAGAGGAAAT 310

Qy 1654 GTATCAGTGTGATCTTTAAGAGGGATGAGACCGGTCAAGCAATGTATACAGTCA 1713
Db 311 GTATCAGTGTGATCTTTAAGAGGGATGAGACCGGTCAAGCAATGTATACAGTCA 370

Qy 1714 TACGAGGAAGCAGCCGATTTATGACTCGAGCCAGATCAATGAAGCATCAAGAGTTTAGG 1773
Db 371 TACGAGGAAGCAGCCGATTTATGACTCGAGCCAGATCAATGAAGCATCAAGAGTTTAGG 430

Qy 1774 AAATGGCATTGGAACCTCAAGAGTACAACGCCAGCGGTGATGACAGTAATCCGACCAT 1833
Db 431 AAATGGCATTGGAACCTCAAGAGTACAACGCCAGCGGTGATGACAGTAATCCGACCAT 490

Qy 1834 GACTATGACTGTACCCGCTCTGGTTCAAGCAGCGAGGGCCAAACACACCGCAATGAG 1893
Db 491 GACTATGACTGTACCCGCTCTGGTTCAAGCAGCGAGGGCCAAACACACCGCAATGAG 550

Qy 1894 ATGGGGAAGATTAAGAGAACCGGTGAGGGTTATAGTGACCTTCTCTTTAA 1944
Db 551 ATGGGGAAGATTAAGAGAACCGGTGAGGGTTATAGTGACCTTCTCTTTAA 601

RESULT 2
BH442220
LOCUS BH442220 BOGLN68TR BOGL Brassica oleracea genomic clone BOGLN68, genomic
DEFINITION survey sequence.
ACCESSION BH442220
VERSION BH442220.1 GI:17627934
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 453)
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Frazer, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BOGLN68TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

REFERENCE
AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Frazer, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOGLN68TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..453
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TOL000DH3"
/db_xref="taxon:3712"
/clone="BOGLN68"
/clone_lib="BOGL"
/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BclXI linkers"

FEATURES
source
1..453
Location/Qualifiers
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TOL000DH3"
/db_xref="taxon:3712"
/clone="BOGLN68"
/clone_lib="BOGL"
/note="Vector: pHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BclXI linkers"

ORIGIN
Query Match      14.0%; Score 273; DB 28; Length 453;
Best Local Similarity 100.0%; Pred. No. 1.6e-112;
Matches 273; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1627 CAGATTGCGTGGCTTAGAAGGAAATGTATCATCTGCAGATCTTAACGAAGGATGAGA 1686
Db 169 CAGATTGCGTGGCTTAGAAGGAAATGTATCATCTGCAGATCTTAACGAAGGATGAGA 228

Qy 1687 CCAGGTCAAGCATGTATACAGCTCATACGAGGAAGCAGCGATTATGACTCGAGCCAG 1746
Db 229 CCAGGTCAAGCATGTATACAGCTCATACGAGGAAGCAGCGATTATGACTCGAGCCAG 288

Qy 1747 TACAATGAAGACATGAAGAGTTTGGAAAATGCACTTGGAACTCAAGAGTACAACGCC 1806
Db 289 TACAATGAAGACATGAAGAGTTTGGAAAATGCACTTGGAACTCAAGAGTACAACGCC 348

Qy 1807 ACGGTCAGTACACTAATCCGACCACTGATGAGCTGATACCCGCTCTGGTTCAAGCAGC 1866
Db 349 ACGGTCAGTACACTAATCCGACCACTGATGAGCTGATACCCGCTCTGGTTCAAGCAGC 408

Qy 1867 GAGGGCCAAACCCACACCGGAATCGAGATGGG 1899
Db 409 GAGGGCCAAACCCACACCGGAATCGAGATGGG 441

RESULT 3
BZ455058
LOCUS BZ455058 BO_1.6.2_KB_tot Brassica oleracea genomic clone BONKA33,
DEFINITION genomic survey sequence.
ACCESSION BZ455058
VERSION BZ455058.1 GI:26731783
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 753)
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Frazer, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other GSSs: BONKA33TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

REFERENCE
AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Frazer, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BONKA33TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..753
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TOL000DH3"
```

```
/db_xref="taxon:3712"
/clone="BONKA33"
/clone_lib="BO 1.6.2 KB tot"
/notes="Vector: pROSL; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pROSL using BstXI linkers"

ORIGIN
Query Match          9.6%; Score 187; DB 28; Length 753;
Best Local Similarity 100.0%; Pred. No. 2e-73;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1282 AACACGATGATCAACACGCTGTGATGGGAACCTTTGGGTACTTGGCTCCGGAAATACGCT 1341
Db |||||
11 AACACGATGATCAACACGCTGTGATGGGAACCTTTGGGTACTTGGCTCCGGAAATACGCT 70
Qy 1342 GCACGCGAAGCTCACGGAAGCTCTGACGCTTTCTCATTTGGCGTGTCTTTGGAG 1401
Db |||||
71 GCAAGCGGAAGCTCACGGAAGCTCTGACGCTTTCTCATTTGGCGTGTCTTTGGAG 130
Qy 1402 CTCATTACTGACGCTGACCGCTTGATGCCAACATGTCTATGTAGATGACAGCTTAGTT 1461
Db |||||
131 CTCATTACTGACGCTGACCGCTTGATGCCAACATGTCTATGTAGATGACAGCTTAGTT 190
Qy 1462 GACTGGG 1468
Db |||||
191 GACTGGG 197

RESULT 4
CD811645
LOCUS      CD811645
DEFINITION BN10.001E08F011207 BN10 Brassica napus cDNA clone BN10001E08, mRNA
sequence.
ACCESSION CD811645
VERSION   CD811645.1 GI:32493585
KEYWORDS EST.
SOURCE    Brassica napus (rape)
ORGANISM  Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 314)
AUTHORS   Genoplatte, a major partnership french program in plant genomics
TITLE     Unpublished (2003)
JOURNAL   Contact: Genoplatte
COMMENT   Genoplatte
          93, rue Henri Rochefort 91025 EVRY CEDEX France
          Tel: 33 1 69 47 54 00
          Fax: 33 1 69 47 54 10
          This sequence has been generated in the framework of the french
          plant genomics programme 'Genoplatte' (http://www.genoplatte.com
          and http://genoplatte-info.infobiogen.fr).
          Location/Qualifiers
            source          1..314
                           /organism="Brassica napus"
                           /mol_type="mRNA"
                           /cultivar="Jet neuf"
                           /db_xref="taxon:3708"
                           /clone="BN1001E08"
                           /tissue_type="seed"
                           /clone_lib="BN10"

ORIGIN
Query Match          3.3%; Score 65; DB 14; Length 314;
Best Local Similarity 99.1%; Pred. No. 3.8e-18;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1825 CCGACCACTGATGACTGACCTGCTACCGCTCTGTTCAAGCAGCGAGGGCCAAACACACGC 1884
Db |||||
1 CCGACCACTGATGACTGACCTGCTGTTTCAAGCAGCGAGGGCCAAACACACGC 60
Qy 1885 GAAATGGAGATGGGGAAGATTAAAGAACCGGTCCAGGTTTATAGTGACCTTCTCT 1940
Db |||||
61 GAGATGGAGATGGGGAAGATTAAAGAACCGGTCCAGGTTTATAGTGACCTTCTCT 116

FEATURES
source
LOCUS      AV539333/c
DEFINITION AV539333 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
cDNA clone RZ130C07F 3', mRNA sequence.
ACCESSION AV539333
VERSION   AV539333.1 GI:8701090
KEYWORDS EST.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 561)
AUTHORS   Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL   DNA Res. 7 (3), 175-180 (2000)
MEDLINE   20363093
```

```
|| |||||
61 GAGATGGAGATGGGGAAGATTAAAGAACCGGTCCAGGTTTATAGTGACCTTCTCT 116

RESULT 5
CD811912
LOCUS      CD811912
DEFINITION BN10.019G03F020121 BN10 Brassica napus cDNA clone BN10019G03, mRNA
sequence.
ACCESSION CD811912
VERSION   CD811912.1 GI:32493852
KEYWORDS EST.
SOURCE    Brassica napus (rape)
ORGANISM  Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 314)
AUTHORS   Genoplatte, a major partnership french program in plant genomics
TITLE     Unpublished (2003)
JOURNAL   Contact: Genoplatte
COMMENT   Genoplatte
          93, rue Henri Rochefort 91025 EVRY CEDEX France
          Tel: 33 1 69 47 54 00
          Fax: 33 1 69 47 54 10
          This sequence has been generated in the framework of the french
          plant genomics programme 'Genoplatte' (http://www.genoplatte.com
          and http://genoplatte-info.infobiogen.fr).
          Location/Qualifiers
            source          1..314
                           /organism="Brassica napus"
                           /mol_type="mRNA"
                           /cultivar="Jet neuf"
                           /db_xref="taxon:3708"
                           /clone="BN10019G03"
                           /tissue_type="seed"
                           /clone_lib="BN10"

ORIGIN
Query Match          3.3%; Score 65; DB 14; Length 314;
Best Local Similarity 99.1%; Pred. No. 3.8e-18;
Matches 115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1825 CCGACCACTGATGACTGACCTGCTACCGCTCTGTTCAAGCAGCGAGGGCCAAACACACGC 1884
Db |||||
1 CCGACCACTGATGACTGACCTGCTGTTTCAAGCAGCGAGGGCCAAACACACGC 60
Qy 1885 GAAATGGAGATGGGGAAGATTAAAGAACCGGTCCAGGTTTATAGTGACCTTCTCT 1940
Db |||||
61 GAGATGGAGATGGGGAAGATTAAAGAACCGGTCCAGGTTTATAGTGACCTTCTCT 116

RESULT 6
AV539333/c
LOCUS      AV539333
DEFINITION AV539333 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
cDNA clone RZ130C07F 3', mRNA sequence.
ACCESSION AV539333
VERSION   AV539333.1 GI:8701090
KEYWORDS EST.
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 561)
AUTHORS   Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL   DNA Res. 7 (3), 175-180 (2000)
MEDLINE   20363093
```

PUBMED 10907847  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.  
 FEATURES  
 source  
 1..561  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="RZ130C07F"  
 /tissue\_type="roots"  
 /clone\_lib="Arabidopsis thaliana roots Columbia"  
 /note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI"  
 ORIGIN  
 Query Match 2.7%; Score 53; DB 9; Length 561;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-12;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1579 GCTGGCGCTTGTTGGCCATTTCAGCTCGCCGACACCTCGCATGAGCCAGAT 1631  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 543 GCTGGCGCTTGTTGGCCATTTCAGCTCGCCGACACCTCGCATGAGCCAGAT 491  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 7  
 AV543493/  
 LOCUS AV543493 Arabidopsis thaliana roots Columbia Arabidopsis thaliana  
 DEFINITION CDNA clone RZ201f07F 3', mRNA sequence.  
 ACCESSION AV543493  
 VERSION AV543493.1 GI:8714907  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 1 (bases 1 to 584)  
 Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.  
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
 of 12,028 non-redundant expressed sequence tags from normalized and  
 size-selected cDNA libraries  
 DNA Res. 7 (3), 175-180 (2000)  
 JOURNAL 20363093  
 MEDLINE 10907847  
 PUBMED  
 COMMENT Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.  
 FEATURES  
 source  
 1..584  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="Columbia"  
 /db\_xref="taxon:3702"  
 /clone="RZ201f07F"  
 /tissue\_type="roots"  
 /clone\_lib="Arabidopsis thaliana roots Columbia"  
 /note="Vector: pBluescriptII SK-; Site\_1: EcoRI; Site\_2: XhoI"  
 ORIGIN  
 Query Match 2.7%; Score 53; DB 9; Length 584;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-12;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1579 GCTGGCGCTTGTTGGCCATTTCAGCTCGCCGACACCTCGCATGAGCCAGAT 1631  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 468 GCTGGCGCTTGTTGGCCATTTCAGCTCGCCGACACCTCGCATGAGCCAGAT 416  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 8  
 BH432371  
 LOCUS BH432371 BOGV Brassica oleracea genomic clone BOGVQ93, genomic  
 DEFINITION survey sequence.  
 ACCESSION BH432371  
 VERSION BH432371.1 GI:17618092  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 754)  
 Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
 Whole genome shotgun sequencing of Brassica oleracea  
 Unpublished (2001)  
 JOURNAL  
 COMMENT Other GSSs: BOGVQ93TF  
 Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..754  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BOGVQ93"  
 /clone\_lib="BOGV"  
 /note="Vector: pHOS1; Site\_1: BstXI; 2-3 kb sheared  
 genomic DNA inserted into pHOS1 using BstXI linkers"  
 ORIGIN  
 Query Match 2.2%; Score 42; DB 28; Length 754;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-07; Mismatches 0; Indels 0; Gaps 0;  
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1762 AAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTACAAC 1803  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 462 AAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTACAAC 503  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 9  
 BZ456327  
 LOCUS BZ456327 BO\_1.6\_2\_KB\_tot Brassica oleracea genomic clone BONBB66,  
 DEFINITION genomic survey sequence.  
 ACCESSION BZ456327  
 VERSION BZ456327.1 GI:26734454  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
 1 (bases 1 to 342)  
 Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
 Whole genome shotgun sequencing of Brassica oleracea  
 Unpublished (2001)  
 JOURNAL  
 COMMENT Contact: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208

Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

# FEATURES

source Location/Qualifiers

1..342  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONBB66"  
/clone\_lib="BO 1.6 2 KB tot"  
/note="Vector: pHS1; Site 1: BstXI; 1.6-2 kb sheared  
total DNA inserted into pHS1 using BstXI linkers"

# ORIGIN

Query Match 1.6%; Score 32; DB 28; Length 342;  
Best Local Similarity 100.0%; Pred. No. 0.0038;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 CCATCTCCACCATCAAACTCCACACACACAC 68

Db 182 CCATCTCCACCATCAAACTCCACACACACAC 213

# RESULT 10

BZ066106/c

LOCUS

DEFINITION BZ066106 744 bp DNA linear GSS 10-OCT-2002  
sequence.

ACCESSION BZ066106

VERSION BZ066106.1

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE 1 (bases 1 to 744)

AUTHORS Dalehaunty,K., Fellw.G., Fulton,L., McCombie,W.R., Miner,T.,

Nash,W., Rabinowicz,P.D. and Wilson,R.K.

TITLE Whole genome shotgun reads from Brassica oleracea

JOURNAL Unpublished (2002)

COMMENT Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: 1j170 row: e column: 12

Seq primer: -28RppOT reverse

Class: shotgun

High quality sequence start: 17

High quality sequence stop: 551.

# FEATURES

source

1..744  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3712"  
/clone\_lib="B.oleracea002"

/note="Vector: pOTw13; Whole genome shotgun library from  
flowering buds. DNA was purified from a crude nuclear  
prep using Brassica oleracea TO1000DH3 buds provided by  
Thomas Osborn at the University of Wisconsin. Genomic  
DNA was provided by Pablo Rabinowicz (CSHL) and the  
shotgun library prepared at Washington University Genome  
Sequencing Center."

# ORIGIN

Query Match 1.5%; Score 30; DB 28; Length 744;  
Best Local Similarity 100.0%; Pred. No. 0.037;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1596 CCATTCAGCTCGCGCAGACCTCGCATGAG 1625

|||||

Db 729 CCATTCAGCTCGCGCAGACCTCGCATGAG 700

# RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

Capiscum annuum

```

COMMENT      Contact: Erika Asamizu
              The First Laboratory for Plant Gene Research
              Kazusa DNA Research Institute
              Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
              Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES     Location/Qualifiers
   source    1..573
              /organism="Arabidopsis thaliana"
              /mol_type="mRNA"
              /strain="Columbia"
              /db_xref="taxon:3702"
              /clone="R2130c07R"
              /tissue_type="roots"
              /clone_lib="Arabidopsis thaliana roots Columbia"
              /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
              XhoI"

ORIGIN
Query Match      1.5%; Score 29; DB 9; Length 573;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      880  TTGCTAGTGGGAAGAGTTGCTGTGAA 908
          |||||||
Db      471  TTGCTAGTGGGAAGAGTTGCTGTGAA 499

RESULT 13
CC968148      607 bp      DNA      linear      GSS 18-AUG-2003
LOCUS      BOIEB10TF BO_1.4_1.6_KB_nuc Brassica oleracea genomic clone
DEFINITION      BOIEB10, genomic survey sequence.
ACCESSION      CC968148
VERSION        CC968148.1 GI:33822075
KEYWORDS       GSS.
SOURCE         Brassica oleracea
ORGANISM       Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 607)
AUTHORS        Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE          Whole genome shotgun sequencing of Brassica oleracea
JOURNAL        Unpublished (2001)
COMMENT        Other GSSs: BOIEB10TR
                Contact: Chris Town
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA.
                Tel: 301-838-3523
                Fax: 301-838-0208
                Email: cdtown@tigr.org
                DNA is from a doubled haploid provided by Tom Osborn.
                Seq primer: TF
                Class: sheared ends.
                Location/Qualifiers
                1..607
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /strain="TO1000DH3"
                /db_xref="taxon:3712"
                /clone="BOIEB10"
                /note="Vector: pPOS2; Site_1: BstXI; 1.4-1.6 kb sheared
                nuclear DNA inserted into pPOS2 using BstXI linkers"

FEATURES     Location/Qualifiers
   source    1..607
              /organism="Brassica oleracea"
              /mol_type="genomic DNA"
              /strain="TO1000DH3"
              /db_xref="taxon:3712"
              /clone="BOIEB10"
              /clone_lib="BO_1.4_1.6_KB_nuc"
              /note="Vector: pPOS2; Site_1: BstXI; 1.4-1.6 kb sheared
              nuclear DNA inserted into pPOS2 using BstXI linkers"

ORIGIN
Query Match      1.5%; Score 29; DB 29; Length 607;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      871  AAAGTGTTGCTAGTGGGAAGAAAGT 899
          |||||||
Db      431  AAAGTGTTGCTAGTGGGAAGAAAGT 459

RESULT 14
BH577077/c    620 bp      DNA      linear      GSS 14-DEC-2001
LOCUS      BOHCX77TR BOHC Brassica oleracea genomic clone BOHCX77, genomic
DEFINITION      survey sequence.
ACCESSION      BH577077
VERSION        BH577077.1 GI:17829358
KEYWORDS       GSS.
SOURCE         Brassica oleracea
ORGANISM       Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 620)
AUTHORS        Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE          Whole genome shotgun sequencing of Brassica oleracea
JOURNAL        Unpublished (2001)
COMMENT        Other GSSs: BOHCX77TF
                Contact: Chris Town
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA.
                Tel: 301-838-3523
                Fax: 301-838-0208
                Email: cdtown@tigr.org
                DNA is from a doubled haploid provided by Tom Osborn.
                Seq primer: TR
                Class: sheared ends.
                Location/Qualifiers
                1..620
                /organism="Brassica oleracea"
                /mol_type="genomic DNA"
                /strain="TO1000DH3"
                /db_xref="taxon:3712"
                /clone="BOHCX77"
                /clone_lib="BOHC"
                /note="Vector: pPOS1; Site_1: BstXI; 2-3 kb sheared
                genomic DNA inserted into pPOS1 using BstXI linkers"

ORIGIN
Query Match      1.5%; Score 29; DB 28; Length 620;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      871  AAAGTGTTGCTAGTGGGAAGAAAGT 899
          |||||||
Db      567  AAAGTGTTGCTAGTGGGAAGAAAGT 539

RESULT 15
BU043317      651 bp      mRNA      linear      EST 26-AUG-2002
LOCUS      PP LEA0015K13f Peach developing fruit mesocarp Prunus persica cDNA
DEFINITION      clone PP_LEA0015K13f, mRNA sequence.
ACCESSION      BU043317
VERSION        BU043317.1 GI:22483394
KEYWORDS       EST.
SOURCE         Prunus persica (peach)
ORGANISM       Prunus persica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
REFERENCE      1 (bases 1 to 651)
AUTHORS        Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.
TITLE          Peach Model Genome for Rosaceae
JOURNAL        Unpublished (2002)
COMMENT        Contact: Abbott, A.
                Dept of Genetics and Biochemistry
                Clemson University
                122 Long Hall, Clemson University, Clemson, SC 29634, USA
                Tel: 864 656 3060
                Fax: 864 656 6879

```

Email: aalbert@clemson.edu  
Total High Quality bases = 578  
Seq primer: TAATACGACTCACTATAGG  
High quality sequence stop: 651.  
FEATURES  
    Location/Qualifiers  
        1..651  
          /organism="Prunus persica"  
          /mol\_type="mRNA"  
          /cultivar="Loring"  
          /db\_xref="taxon:3760"  
          /clone="PP\_LEA0015K13f"  
          /tissue\_type="Mesocarp"  
          /lab\_host="E. coli"  
          /clone\_lib="Peach developing fruit mesocarp"  
          /note="Vector: pBluescript II SK(-); Site 1: EcoRI;  
          Site 2: XhoI; authority=Prunus persica L. Batsh; The  
          sequence has been trimmed to remove vector sequence and  
          contains a minimum of 100 bases of phred value 20 or  
          above. For more details on library preparation and  
          sequence analysis go to  
          http://www.genome.clemson.edu/projects/peach. To order  
          this clone go to http://www.genome.clemson.edu/orders"

ORIGIN

Query Match 1.5%; Score 29; DB 13; Length 651;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1555 GAGATGGCTCGCATGGTTGCTTGCTGCTGC 1583  
      |||||  
Db 270 GAGATGGCTCGCATGGTTGCTTGCTGCTGC 298

Search completed: May 12, 2004, 20:46:18  
Job time : 4562 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2004, 16:54:59 ; Search time 168 Seconds  
(without alignments)  
6421.575 Million cell updates/sec

Title: US-10-086-464-1  
Perfect score: 1944  
Sequence: 1 atgtcctggcgccgtctcc.....atagtgacattctcttaa 1944

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	1.2	4626	US-08-306-691B-22	Sequence 22, Appl
2	23	1.2	4626	PCT-US93-06251-27	Sequence 27, Appl
3	20	1.0	609	US-09-800-729-55	Sequence 55, Appl
4	20	1.0	852	US-09-252-991A-1274	Sequence 1274, Ap
5	20	1.0	909	US-09-252-991A-1381	Sequence 1381, Ap
6	20	1.0	1092	US-09-252-991A-1287	Sequence 1287, Ap
7	20	1.0	3097	US-09-228-986-11	Sequence 11, Appl
8	20	1.0	5720	US-09-800-729-18	Sequence 18, Appl
9	20	1.0	41100	US-09-755-665-46	Sequence 46, Appl
10	19	1.0	358	US-08-991-789A-168	Sequence 168, App
11	19	1.0	358	US-09-062-451-168	Sequence 168, App
12	19	1.0	358	US-09-598-326-168	Sequence 168, App
13	19	1.0	358	US-09-289-198-168	Sequence 168, App
14	19	1.0	358	US-09-429-755-168	Sequence 168, App
15	19	1.0	2421	US-09-187-330-4	Sequence 4, Appl
16	19	1.0	2487	US-09-187-330-54	Sequence 54, Appl
17	19	1.0	3621	US-09-635-872A-21	Sequence 21, Appl
18	19	1.0	3621	US-09-636-077A-21	Sequence 21, Appl
19	19	1.0	3621	US-09-636-060C-21	Sequence 21, Appl
20	19	1.0	3621	US-09-986-552-21	Sequence 21, Appl
21	19	1.0	3783	US-09-635-872A-20	Sequence 20, Appl
22	19	1.0	3783	US-09-636-077A-20	Sequence 20, Appl
23	19	1.0	3783	US-09-636-060C-20	Sequence 20, Appl
24	19	1.0	3783	US-09-986-552-20	Sequence 20, Appl
25	19	1.0	5053	US-09-376-330-1	Sequence 1, Appl
26	19	1.0	5597	US-09-635-872A-4	Sequence 4, Appl
27	19	1.0	5597	US-09-636-077A-4	Sequence 4, Appl

c 28	19	1.0	5597	4	US-09-636-060C-4	Sequence 4, Appl
c 29	19	1.0	5597	4	US-09-986-552-4	Sequence 4, Appl
c 30	18	0.9	274	4	US-09-313-294A-2260	Sequence 2260, Ap
c 31	18	0.9	293	4	US-09-313-294A-6390	Sequence 6390, Ap
c 32	18	0.9	300	4	US-09-313-294A-4708	Sequence 4708, Ap
c 33	18	0.9	306	4	US-09-313-294A-4079	Sequence 4079, Ap
c 34	18	0.9	308	4	US-09-313-294A-5000	Sequence 5000, Ap
c 35	18	0.9	424	4	US-09-621-976-9270	Sequence 9270, Ap
c 36	18	0.9	439	4	US-09-397-787-273	Sequence 273, App
c 37	18	0.9	489	1	US-07-879-685B-3	Sequence 3, Appl
c 38	18	0.9	652	4	US-09-976-594-666	Sequence 666, App
c 39	18	0.9	657	4	US-09-833-381-921	Sequence 921, App
c 40	18	0.9	684	3	US-08-998-416-672	Sequence 672, App
c 41	18	0.9	1008	4	US-09-500-569-1	Sequence 1, Appl
c 42	18	0.9	1008	4	US-09-971-823B-1	Sequence 1, Appl
c 43	18	0.9	1276	3	US-09-177-325-2	Sequence 2, Appl
c 44	18	0.9	1276	3	US-09-411-812A-2	Sequence 2, Appl
c 45	18	0.9	1276	4	US-09-590-113-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-306-691B-22  
; Sequence 22, Application US/08306691B  
; Patent No. 5734039  
; GENERAL INFORMATION:  
; APPLICANT: Calabretta, Bruno  
; APPLICANT: Skorski, Tomasz  
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Seidel, Gonda, Lavorgha & Monaco, P.C.  
; STREET: Two Penn Center, Suite 1800  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306,691B  
; FILING DATE: September 15, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 8321-8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; TELEX: No. 5734039e  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4626 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-306-691B-22

Query Match 1-2%; Score 23; DB 1; Length 4626;  
Best Local Similarity 100.0%; Pred. No. 0.27; Indels 0; Gaps 0;  
Matches 23; Conservative 0; Mismatches 0;  
Oy 1240 AAGTGTGCTGATTGTTGGTCTTGC 1262

```
Db          3903 AAGGTGCTGATTTGGTCTTGC 3925
|||||
PCT-US93-06251-27
; Sequence 27, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4626 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-27

Query Match          1.2%; Score 23; DB 5; Length 4626;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1240 AAGGTGCTGATTTGGTCTTGC 1262
|||||
Db          3903 AAGGTGCTGATTTGGTCTTGC 3925
|||||

RESULT 3
US-09-800-729-55
; Sequence 55, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: PZ044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 609
; TYPE: DNA

Query Match          1.2%; Score 23; DB 5; Length 4626;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (554)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-800-729-55

Query Match          1.0%; Score 20; DB 4; Length 609;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          667 CAACCTCTCCACCTCCACC 686
|||||
Db          269 CAACCTCTCCACCTCCACC 288
|||||

RESULT 4
US-09-252-991A-1274
; Sequence 1274, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1274
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1274

Query Match          1.0%; Score 20; DB 4; Length 852;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY          1467 GGCACGACCATTTGCTTAACC 1486
|||||
Db          56 GGCACGACCATTTGCTTAACC 75
|||||

RESULT 5
US-09-252-991A-1381/c
; Sequence 1381, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1381
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1381

Query Match          1.0%; Score 20; DB 4; Length 909;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1467 GGCACGACCATGCTTAACC 1486  
Db 112 GGCACGACCATGCTTAACC 93

## RESULT 6

US-09-252-991A-1287/c  
; Sequence 1287, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 1287  
; LENGTH: 1092  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-1287

Query Match 1.0%; Score 20; DB 4; Length 1092;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1467 GGCACGACCATGCTTAACC 1486  
Db 126 GGCACGACCATGCTTAACC 107

## RESULT 7

US-09-228-986-11  
; Sequence 11, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 1100/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 11  
; LENGTH: 3097  
; TYPE: DNA  
; ORGANISM: Eucalyptus grandis  
US-09-228-986-11

Query Match 1.0%; Score 20; DB 4; Length 3097;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1387 GTTGCTCTTTGGAGCTCAT 1406  
Db 2556 GTTGCTCTTTGGAGCTCAT 2575

## RESULT 8

US-09-800-729-18  
; Sequence 18, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1

; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 5720  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-800-729-18

Query Match 1.0%; Score 20; DB 4; Length 5720;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 667 CAACCTCTCCACCTCCACC 686  
Db 5431 CAACCTCTCCACCTCCACC 5450

## RESULT 9

US-09-755-665-46/c  
; Sequence 46, Application US/09755665  
; Patent No. 6600019  
; GENERAL INFORMATION:  
; APPLICANT: Prayaga, Sudhirdas K.  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Tailon, Bruce E.  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Spvtek, Kimberly A.  
; APPLICANT: MacDougall, John  
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15966-631  
; CURRENT APPLICATION NUMBER: US/09/755,665  
; CURRENT FILING DATE: 2001-08-14  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/174,724  
; PRIOR FILING DATE: 2000-01-06  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 46  
; LENGTH: 41100  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: GENOMIC DNA  
US-09-755-665-46

Query Match 1.0%; Score 20; DB 4; Length 41100;  
Best Local Similarity 100.0%; Pred. No. 7.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 921 TGGGAGTGGTCAGGAGAGA 940  
Db 22048 TGGGAGTGGTCAGGAGAGA 22029

## RESULT 10

US-08-991-789A-168  
; Sequence 168, Application US/08991789A  
; Patent No. 6225054  
; GENERAL INFORMATION:  
; APPLICANT: Fridakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; NUMBER OF SEQUENCES: 292  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed IP Law Group  
; STREET: 701 Fifth Avenue, Suite 6300

```
;
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 168:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 358 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-08-991-789A-168

Query Match          1.0%; Score 19; DB 3; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      950 AGGCAGAGGTTGAGATCAT 968
Db      228 AGGCAGAGGTTGAGATCAT 246

RESULT 11
US-09-062-451-168
; Sequence 168, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 237
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 168:
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 358 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-062-451-168

Query Match          1.0%; Score 19; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      950 AGGCAGAGGTTGAGATCAT 968
Db      228 AGGCAGAGGTTGAGATCAT 246

RESULT 12
US-09-598-326-168
; Sequence 168, Application US/09598326
; Patent No. 6423496
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 247
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group PLLC
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/598,326
; FILING DATE: 20-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E.R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 168:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 358 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 168:
US-09-598-326-168

Query Match          1.0%; Score 19; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      950 AGGCAGAGGTTGAGATCAT 968
Db      228 AGGCAGAGGTTGAGATCAT 246

RESULT 13
US-09-289-198-168
; Sequence 168, Application US/09289198
; Patent No. 6586570
; GENERAL INFORMATION:
```

; APPLICANT: Frudakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Misher, Lynda  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.419C5  
; CURRENT APPLICATION NUMBER: US/09/289,198  
; CURRENT FILING DATE: 1999-04-09  
; EARLIER APPLICATION NUMBER: US 09/062,451  
; EARLIER FILING DATE: 1998-04-17  
; EARLIER APPLICATION NUMBER: US 08/991,789  
; EARLIER FILING DATE: 1997-12-11  
; EARLIER APPLICATION NUMBER: US 08/838,762  
; EARLIER FILING DATE: 1997-04-09  
; EARLIER APPLICATION NUMBER: PCT/US97/00485  
; EARLIER FILING DATE: 1997-01-10  
; EARLIER APPLICATION NUMBER: US 08/700,014  
; EARLIER FILING DATE: 1996-08-20  
; EARLIER APPLICATION NUMBER: US 08/585,392  
; EARLIER FILING DATE: 1996-01-01  
; NUMBER OF SEQ ID NOS: 312  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 168  
; LENGTH: 358  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-289-198-168

Query Match 1.0%; Score 19; DB 4; Length 358;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 950 AGGCAGAGGTTGAGATCAT 968  
Db 228 AGGCAGAGGTTGAGATCAT 246  
|||||

## RESULT 14

US-09-429-755-168  
; Sequence 168, Application US/09429755A  
; Patent No. 6656480  
; GENERAL INFORMATION:  
; APPLICANT: Frudakis, Tony N.  
; APPLICANT: Smith, John M.  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Misher, Lynda  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.419C6  
; CURRENT APPLICATION NUMBER: US/09/429,755A  
; CURRENT FILING DATE: 1999-10-28  
; NUMBER OF SEQ ID NOS: 315  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 168  
; LENGTH: 358  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-429-755-168

Query Match 1.0%; Score 19; DB 4; Length 358;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 950 AGGCAGAGGTTGAGATCAT 968  
Db 228 AGGCAGAGGTTGAGATCAT 246  
|||||

## RESULT 15

US-09-187-330-4/c

; Sequence 4, Application US/09187330  
; Patent No. 6613740  
; GENERAL INFORMATION:  
; APPLICANT: Gozes, Illana  
; APPLICANT: Brennenman, Douglas E.  
; APPLICANT: Bassan, Merav  
; APPLICANT: Zamostiano, Rachel  
; APPLICANT: The Government of the United States of America  
; APPLICANT: as represented by the Secretary of the  
; APPLICANT: Department of Health and Human Services  
; TITLE OF INVENTION: Activity Dependent Neurotrophic Factor III (ADNF III)  
; FILE REFERENCE: 015280-291200US  
; CURRENT APPLICATION NUMBER: US/09/187,330  
; CURRENT FILING DATE: 1998-11-06  
; EARLIER APPLICATION NUMBER: US 60/037,404  
; EARLIER FILING DATE: 1997-02-07  
; EARLIER APPLICATION NUMBER: WO PCT/US98/02485  
; EARLIER FILING DATE: 1998-02-06  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 2421  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; OTHER INFORMATION: one mouse activity dependent neurotrophic factor  
; OTHER INFORMATION: III (ADNF III) cDNA clone  
US-09-187-330-4

Query Match 1.0%; Score 19; DB 4; Length 2421;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 TCCTCCTCCAGCTTCGCGCT 87  
Db 1295 TCCTCCTCCAGCTTCGCGCT 1277  
|||||

Search completed: May 12, 2004, 20:49:09  
Job time : 173 secs

**This Page Blank (uspto)**

Result No.	Score	Query		ID	Description
		Match	Length		
1	1944	100.0	1944	14	US-10-086-464-1
2	1944	100.0	2189	14	US-10-086-464-3
3	26	1.3	268	9	US-09-923-876-5096
4	26	1.3	268	11	US-09-923-876-5096
5	26	1.3	951	13	US-10-425-114-1922
6	26	1.3	1011	13	US-10-425-114-13131
7	26	1.3	1134	13	US-10-425-114-36420
8	26	1.3	1500	13	US-10-425-114-6300
9	26	1.3	1552	13	US-10-424-599-47702
10	26	1.3	1724	13	US-10-425-114-2326
11	25	1.3	1101	13	US-10-424-599-122411
12	24	1.2	1515	13	US-10-368-046-1
13	24	1.2	1515	16	US-10-367-095-1
14	23	1.2	180	15	US-10-029-386-26494
15	23	1.2	180	15	US-10-029-386-26494
16	23	1.2	180	15	US-10-029-386-26494
17	23	1.2	180	15	US-10-029-386-26494
18	23	1.2	180	15	US-10-029-386-26494
19	23	1.2	180	15	US-10-029-386-26494
20	23	1.2	180	15	US-10-029-386-26494
21	23	1.2	180	15	US-10-029-386-26494
22	23	1.2	180	15	US-10-029-386-26494
23	23	1.2	180	15	US-10-029-386-26494
24	23	1.2	180	15	US-10-029-386-26494
25	23	1.2	180	15	US-10-029-386-26494
26	23	1.2	180	15	US-10-029-386-26494
27	23	1.2	180	15	US-10-029-386-26494
28	23	1.2	180	15	US-10-029-386-26494
29	23	1.2	180	15	US-10-029-386-26494
30	23	1.2	180	15	US-10-029-386-26494
31	23	1.2	180	15	US-10-029-386-26494
32	23	1.2	180	15	US-10-029-386-26494
33	23	1.2	180	15	US-10-029-386-26494
34	23	1.2	180	15	US-10-029-386-26494
35	23	1.2	180	15	US-10-029-386-26494
36	23	1.2	180	15	US-10-029-386-26494
37	23	1.2	180	15	US-10-029-386-26494
38	23	1.2	180	15	US-10-029-386-26494
39	23	1.2	180	15	US-10-029-386-26494
40	23	1.2	180	15	US-10-029-386-26494
41	23	1.2	180	15	US-10-029-386-26494
42	23	1.2	180	15	US-10-029-386-26494
43	23	1.2	180	15	US-10-029-386-26494
44	23	1.2	180	15	US-10-029-386-26494
45	23	1.2	180	15	US-10-029-386-26494
46	23	1.2	180	15	US-10-029-386-26494
47	23	1.2	180	15	US-10-029-386-26494
48	23	1.2	180	15	US-10-029-386-26494
49	23	1.2	180	15	US-10-029-386-26494
50	23	1.2	180	15	US-10-029-386-26494
51	23	1.2	180	15	US-10-029-386-26494
52	23	1.2	180	15	US-10-029-386-26494
53	23	1.2	180	15	US-10-029-386-26494
54	23	1.2	180	15	US-10-029-386-26494
55	23	1.2	180	15	US-10-029-386-26494
56	23	1.2	180	15	US-10-029-386-26494
57	23	1.2	180	15	US-10-029-386-26494
58	23	1.2	180	15	US-10-029-386-26494
59	23	1.2	180	15	US-10-029-386-26494
60	23	1.2	180	15	US-10-029-386-26494
61	23	1.2	180	15	US-10-029-386-26494
62	23	1.2	180	15	US-10-029-386-26494
63	23	1.2	180	15	US-10-029-386-26494

QY 61 ACCACCACTCTCTCCAGCTTCGGCTCCTCTCCACACACCTTCTTCTCTCCGCG 120  
Db 61 ACCACCACTCTCTCCAGCTTCGGCTCCTCTCCACACACCTTCTTCTCTCCGCG 120  
QY 121 CCATCCACTATTTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
Db 121 CCATCCACTATTTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
QY 181 CCATCTCCACCAACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCA 240  
Db 181 CCATCTCCACCAACTCCATCTACGCGGGATCTCCACCTCTCTCTCTCTCTCTCTCTCA 240  
QY 241 CCGCTCCAACTACGCGCGGATCTCCACCGGACCTGTACTCTCTCTCTCTCTCTCTCTCA 300  
Db 241 CCGCTCCAACTACGCGCGGATCTCCACCGGACCTGTACTCTCTCTCTCTCTCTCTCTCA 300  
QY 301 CCACCTTCAGTCCGAGGACCGGTCCTCAATCTCTCCGCGGAGGAGGATCTCTCGACCT 360  
Db 301 CCACCTTCAGTCCGAGGACCGGTCCTCAATCTCTCCGCGGAGGAGGATCTCTCGACCT 360  
QY 361 CCATCTTCTCCTCGCGCGCT 420  
Db 361 CCATCTTCTCCTCGCGCGCT 420  
QY 421 GGAATCGCATCGGAGGATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
Db 421 GGAATCGCATCGGAGGATCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480  
QY 481 AAGAAGAAACGAGGAGAGAGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
Db 481 AAGAAGAAACGAGGAGAGAGAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
QY 541 GGTCCCAAAGCGGAGGACCTTACGGTGGACAGCAACAATGGCGGCAACAAACGCA 600  
Db 541 GGTCCCAAAGCGGAGGACCTTACGGTGGACAGCAACAATGGCGGCAACAAACGCA 600  
QY 601 ACACACCGTCAGATCATGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 660  
Db 601 ACACACCGTCAGATCATGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 660  
QY 661 CCACGGCAACCTCTCTCCACCTCCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720  
Db 661 CCACGGCAACCTCTCTCCACCTCCACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720  
QY 721 TACTCGGACCGTCCAGTCT 780  
Db 721 TACTCGGACCGTCCAGTCT 780  
QY 781 AGCACTTTTCATACGAGGAGTAGCTAGAGCCCAATGGTCTCTCTCTCTCTCTCTCTCT 840  
Db 781 AGCACTTTTCATACGAGGAGTAGCTAGAGCCCAATGGTCTCTCTCTCTCTCTCTCTCT 840  
QY 841 TTAGGACAAGCGGGTTCGGTTACGTGCACAAAGTGTGTCTCTCTCTCTCTCTCTCTCT 900  
Db 841 TTAGGACAAGCGGGTTCGGTTACGTGCACAAAGTGTGTCTCTCTCTCTCTCTCTCTCT 900  
QY 901 GCTGTGAAGCAGTTGAAAGTTGGGAGTGTGAGGAGAGGAGTTCAGGAGAGTTCAGGAGT 960  
Db 901 GCTGTGAAGCAGTTGAAAGTTGGGAGTGTGAGGAGAGGAGTTCAGGAGAGTTCAGGAGT 960  
QY 961 GAGATCATCAGAGATTCACACAGGATCTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020  
Db 961 GAGATCATCAGAGATTCACACAGGATCTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020  
QY 1021 GGTGCCAAAAGATTCCTGTCTATCAGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080  
Db 1021 GGTGCCAAAAGATTCCTGTCTATGAGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080  
QY 1081 CATGCGGAGGAGCGGCTCAATGGAATGGAGCACCAGATTCGAAGTTGCTCTCTCTCTCT 1140  
Db 1081 CATGCGGAGGAGCGGCTCAATGGAATGGAGCACCAGATTCGAAGTTGCTCTCTCTCTCT 1140  
QY 1141 GCTAAAGGACCTTCTTATCTTCATGAAGTTGCAATCTCTCTCTCTCTCTCTCTCTCTCT 1200

Db 1141 GCTAAAGGACCTTCTTATCTTCATGAAGTTGCAATCTCTCTCTCTCTCTCTCTCTCTCT 1200  
QY 1201 AAGGCTTCAACACATATTTAGTAGATTTCAAGTTTGAAGCTTAAGTTTGTCTTATTTGGTCT 1260  
Db 1201 AAGGCTTCAACACATATTTAGTAGATTTCAAGTTTGAAGCTTAAGTTTGTCTTATTTGGTCT 1260  
QY 1261 GCTAAGATTGCTTCTGATCAAAACGCGATGTATCAACGCTGTGATGGGAACCTTTGGG 1320  
Db 1261 GCTAAGATTGCTTCTGATCAAAACGCGATGTATCAACGCTGTGATGGGAACCTTTGGG 1320  
QY 1321 TACTTGGCTCGGAATAGCTGCAAGCGGAAGCTCAAGGAAAGTCTGAGCTTTTCTCA 1380  
Db 1321 TACTTGGCTCGGAATAGCTGCAAGCGGAAGCTCAAGGAAAGTCTGAGCTTTTCTCA 1380  
QY 1381 TTTGGCGTGTGCTTTTGGAGCTCATTACTGAGAGTCCAGCCGTTGATGCCAACAAATGTC 1440  
Db 1381 TTTGGCGTGTGCTTTTGGAGCTCATTACTGAGAGTCCAGCCGTTGATGCCAACAAATGTC 1440  
QY 1441 TATGTAGATGACAGCTTAGTTGACTGGGACGACATTTGCTTAACCGAGGATCTGAGCAA 1500  
Db 1441 TATGTAGATGACAGCTTAGTTGACTGGGACGACATTTGCTTAACCGAGGATCTGAGCAA 1500  
QY 1501 GGAGACTTTGAGGGTGTAGCTGATGCAAGATGAATATGAGGATGACAGAGAGAGATG 1560  
Db 1501 GGAGACTTTGAGGGTGTAGCTGATGCAAGATGAATATGAGGATGACAGAGAGAGATG 1560  
QY 1561 GCTCGCATGTTGCTTGTGCTCGGCTTGTGCTCGCATTCAGCTCGCGCAGACCTCGC 1620  
Db 1561 GCTCGCATGTTGCTTGTGCTCGGCTTGTGCTCGCATTCAGCTCGCGCAGACCTCGC 1620  
QY 1621 ATGAGCCAGATTGCTGCTGCTTGAAGAAATGTATCAGTCTGATGATCTTAAACGAGGG 1680  
Db 1621 ATGAGCCAGATTGCTGCTGCTTGAAGAAATGTATCAGTCTGATGATCTTAAACGAGGG 1680  
QY 1681 ATGAGCCAGCTCAAGCAATGTATACAGCTCATACGAGGAGAACCCAGATTATGACTCG 1740  
Db 1681 ATGAGCCAGCTCAAGCAATGTATACAGCTCATACGAGGAGAACCCAGATTATGACTCG 1740  
QY 1741 AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800  
Db 1741 AGCCAGTACATGAAGACATGAAGAAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800  
QY 1801 AACGCCAGGTTGAGTACAGTAATCCGACGAGTACATGAGCTGTACCCGTCTGGTTCA 1860  
Db 1801 AACGCCAGGTTGAGTACAGTAATCCGACGAGTACATGAGCTGTACCCGTCTGGTTCA 1860  
QY 1861 AGCAGCGAGGCGCAAAACACACGCAATGAGATGGGAAAGATTAAAGAAACCGGTGAG 1920  
Db 1861 AGCAGCGAGGCGCAAAACACACGCAATGAGATGGGAAAGATTAAAGAAACCGGTGAG 1920  
QY 1921 GGTATAGTGACCTTCTCTTTAA 1944  
Db 1921 GGTATAGTGACCTTCTCTTTAA 1944

## RESULT 2

US-10-086-464-3  
; Sequence 3, Application US/10086464  
; Publication No. US20020199218A1  
; GENERAL INFORMATION:  
; APPLICANT: GORING, Daphne R. et al.  
; TITLE OF INVENTION: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES  
; FILE REFERENCE: P 25,762-A USA  
; CURRENT APPLICATION NUMBER: US/10/086,464  
; FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 10/069,304  
; FILING DATE: 2002-02-19  
; PRIOR APPLICATION NUMBER: PCT/CA00/00966  
; FILING DATE: 2000-08-18  
; PRIOR APPLICATION NUMBER: US 60/149,466  
; FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: US 60/159,122

```
; PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2189
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2189)
; OTHER INFORMATION:
US-10-086-464-3

Query Match      100.0%; Score 1944; DB 14; Length 2189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1944; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTCTCGGCGCGCTCTCGGGGAGCTGGTTGGCTCCATCTCCACCACTCAAACTCCACA 60
Db 97 ATGCTCTCGGCGCGCTCTCGGGGAGCTGGTTGGCTCCATCTCCACCACTCAAACTCCACA 156

Qy 61 ACCACCACTCTCTCCAGCTCTCGCTCTCTCCACCACTCTCTCTCTCTCTCTCTCTCTCT 120
Db 157 ACCACCACTCTCTCTCCAGCTCTCGCTCTCTCCACCACTCTCTCTCTCTCTCTCTCTCTCT 216

Qy 121 CCATCCACTATTTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 217 CCATCCACTATTTCCGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 276

Qy 181 CCATCTCCAACTCCATCTAGCGCGGGATCTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Db 277 CCATCTCCAACTCCATCTAGCGCGGGATCTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 336

Qy 241 CCCGCTCAACTAGCGCGGGATCTCTCCACCGCACTGTACTCTCTCTCTCTCTCTCTCTCTCT 300
Db 337 CCCGCTCAACTAGCGCGGGATCTCTCCACCGCACTGTACTCTCTCTCTCTCTCTCTCTCTCT 396

Qy 301 CCACCTTCAGTCCAGAGACACCGTCCAACTCTTCCAGGAGGAGATCTCTCTCGACCT 360
Db 397 CCACCTTCAGTCCAGAGACACCGTCCAACTCTTCCAGGAGGAGATCTCTCTCGACCT 456

Qy 361 CCATCTCTCTCTCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
Db 457 CCATCTCTCTCTCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 516

Qy 421 GGAATCGGCATCGGAGGAGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Db 517 GGAATCGGCATCGGAGGAGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576

Qy 481 AAGAAGAAACGACGAGAGAGAGAGAGATGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCT 540
Db 577 AAGAAGAAACGACGAGAGAGAGAGAGATGCTTACTATGTTCTCTCTCTCTCTCTCTCTCTCT 636

Qy 541 GGTCCCAAGCGGAGGAGCTTTACGGTGGACAGACGACCACTTCCGCGCAACAAACGCA 600
Db 637 GGTCCCAAGCGGAGGAGCTTTACGGTGGACAGACGACCACTTCCGCGCAACAAACGCA 696

Qy 601 ACACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 697 ACACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 756

Qy 661 CCAGGCAACCTCTCTCCACCTCCACCGGCTTCTATGAGCAGCAGCGGCGGCTCCGAC 720
Db 757 CCAGGCAACCTCTCTCCACCTCCACCGGCTTCTATGAGCAGCAGCGGCGGCTCCGAC 816

Qy 721 TACTCGGACCGTCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 817 TACTCGGACCGTCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 876

Qy 781 AGCACTTTTCATACGAGGAGCTAGCTAGAGCACCCTTCTCTCTCTCTCTCTCTCTCTCTCT 840
Db 877 AGCACTTTTCATACGAGGAGCTAGCTAGAGCACCCTTCTCTCTCTCTCTCTCTCTCTCTCT 936
```

```
Qy 841 TTAGCACAAGCGCGGTTCCGTTACGTGCACAAAGGTGTGTGTCCTAGTGGGAAAGAGTT 900
Db 937 TTAGCACAAGCGCGGTTCCGTTACGTGCACAAAGGTGTGTGTCCTAGTGGGAAAGAGTT 996

Qy 901 GCTGTGAAGCAGTTGAAGTTGGAGTGGTTCAGGAGAGAGGAGTTTCAGGAGAGAGTT 960
Db 997 GCTGTGAAGCAGTTGAAGTTGGAGTGGTTCAGGAGAGAGGAGTTTCAGGAGAGAGTT 1056

Qy 961 GAGATCATCAGCAGAGTTCCACCAAGCATCTGCTGTCTCTTGTGGTATTTGATCGCC 1020
Db 1057 GAGATCATCAGCAGAGTTCCACCAAGCATCTGCTGTCTCTTGTGGTATTTGATCGCC 1116

Qy 1021 GGTCCAAAAGATTTGCTGTCTATGATTTGTTCTTAAACAATCTCGAGCTTCACCTC 1080
Db 1117 GGTCCAAAAGATTTGCTGTCTATGATTTGTTCTTAAACAATCTCGAGCTTCACCTC 1176

Qy 1081 CATGCCAGGAGCGGCTTACATGGAATGAGCAGCAGATTTGAAGATTTGCTTTGGATCT 1140
Db 1177 CATGCCAGGAGCGGCTTACATGGAATGAGCAGCAGATTTGAAGATTTGCTTTGGATCT 1236

Qy 1141 GCTAAAGGACTTTCTTATCTTCTATGAAGATTTGCAATCTTAAATCATTCACCGTATATC 1200
Db 1237 GCTAAAGGACTTTCTTATCTTCTATGAAGATTTGCAATCTTAAATCATTCACCGTATATC 1296

Qy 1201 AAGCTTTCAACATATTTGATAGATTTCAAGTTTGAAGCTAAGTTGCTGATTTGGTCTT 1260
Db 1297 AAGCTTTCAACATATTTGATAGATTTCAAGTTTGAAGCTAAGTTGCTGATTTGGTCTT 1356

Qy 1261 GCTAAGATTTGCTTCTGATACAAACAGCATGTATCAACAGCTGTGATGGGAACCTTTGG 1320
Db 1357 GCTAAGATTTGCTTCTGATACAAACAGCATGTATCAACAGCTGTGATGGGAACCTTTGG 1416

Qy 1321 TACTTGTCTCGGAATACGCTCAAGCGGAAGCTCAGCGGAAGTCTCGAGTCTTCTCTCA 1380
Db 1417 TACTTGTCTCGGAATACGCTCAAGCGGAAGCTCAGCGGAAGTCTCGAGTCTTCTCTCA 1476

Qy 1381 TTTGGCGTGTGCTTTTGGAGCTCATTA CTGGAGCTCACTGAGCGTCCAGCCGTTGATGCAACAATGTC 1440
Db 1477 TTTGGCGTGTGCTTTTGGAGCTCATTA CTGGAGCTCACTGAGCGTCCAGCCGTTGATGCAACAATGTC 1536

Qy 1441 TATGTAGATGACAGCTTAGTTGACTGGGCGACGACCATTTGCTTAACCGAGCATCTGAGCAA 1500
Db 1537 TATGTAGATGACAGCTTAGTTGACTGGGCGACGACCATTTGCTTAACCGAGCATCTGAGCAA 1596

Qy 1501 GGACACTTTGAGGGTTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Db 1597 GGACACTTTGAGGGTTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1656

Qy 1561 GCTCGCATGGTGTGCTGTGCTGGGCTTGTGTTCCCAATTCAGCTCGCGCGCAGACCTCGC 1620
Db 1657 GCTCGCATGGTGTGCTGTGCTGGGCTTGTGTTCCCAATTCAGCTCGCGCGCAGACCTCGC 1716

Qy 1621 ATGAGCAGATTTGCTGTGCTGCTTGAAGGAAATGTATCATCTGTGATCTTTAACAAGGG 1680
Db 1717 ATGAGCAGATTTGCTGTGCTGCTTGAAGGAAATGTATCATCTGTGATCTTTAACAAGGG 1776

Qy 1681 ATGAGCAGCTCAAGCAATGTATACAGCTCATACGAGGAAAGCAGATTTATGACTCG 1740
Db 1777 ATGAGCAGCTCAAGCAATGTATACAGCTCATACGAGGAAAGCAGATTTATGACTCG 1836

Qy 1741 AGCCAGTCAATGAAGACATGAAGAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1800
Db 1837 AGCCAGTCAATGAAGACATGAAGAGTTTAGGAAAATGGCACTTGGAACTCAAGAGTAC 1896

Qy 1801 AACGCCACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1897 AACGCCACGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1956

Qy 1861 AGCAGCGGCGGCAACCAACAGCGGAAATGAGATGGGAAAGATTTAGAGAACCGGTGAG 1920
Db 1957 AGCAGCGGCGGCAACCAACAGCGGAAATGAGATGGGAAAGATTTAGAGAACCGGTGAG 2016

Qy 1921 GGTTATAGTGAGCCTTCTCTTTAA 1944
```

```
|||||
Db      2017 GATTATAGTGACCTTCTCTTAA 2040

RESULT 3
US-09-923-876-5096
; Sequence 5096, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5096
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc_feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700456354H1
; NAME/KEY: unsure
; LOCATION: 63
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5096

Query Match      1.3%; Score 26; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1231 TTGAAGCTAAGTTGCTGATTTGG 1256
      |||||||
Db      203 TTGAAGCTAAGTTGCTGATTTGG 228

RESULT 4
US-09-923-876-5096
; Sequence 5096, Application US/09923876
; Publication No. US20030237110A9
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5096
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc_feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20030237110A9 700456354H1
; NAME/KEY: unsure
; LOCATION: 63
; OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-5096

Query Match      1.3%; Score 26; DB 9; Length 268;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1231 TTGAAGCTAAGTTGCTGATTTGG 1256
      |||||||
Db      203 TTGAAGCTAAGTTGCTGATTTGG 228

RESULT 5
US-10-425-114-1922
; Sequence 1922, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1922
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700171392_FLI
US-10-425-114-1922

Query Match      1.3%; Score 26; DB 13; Length 951;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1231 TTGAAGCTAAGTTGCTGATTTGG 1256
      |||||||
Db      16 TTGAAGCTAAGTTGCTGATTTGG 41

RESULT 6
US-10-425-114-13131
; Sequence 13131, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13131
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMFL02220096D10_FLI
US-10-425-114-13131

Query Match      1.3%; Score 26; DB 13; Length 1011;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1360 GAGAAGCTGACGTTTCTCATTGG 1385
```

Db 246 GAGAAGTCTGACGTTTCTCATTTGG 271  
|||||

RESULT 7  
US-10-425-114-36420  
; Sequence 36420, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 36420  
; LENGTH: 1134  
; TYPE: DNA  
; ORGANISM: Zea mays subsp. mexicana  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMROTBOSINTE097B02\_FLI  
US-10-425-114-36420

Query Match 1.3%; Score 26; DB 13; Length 1134;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1231 TTGAAGCTAGGTTGCTGATTTGG 1256  
|||||  
Db 260 TTGAAGCTAGGTTGCTGATTTGG 285  
|||||

RESULT 8  
US-10-425-114-6300  
; Sequence 6300, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 6300  
; LENGTH: 1500  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700560157\_FLI  
US-10-425-114-6300

Query Match 1.3%; Score 26; DB 13; Length 1500;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1360 GAGAAGTCTGACGTTTCTCATTTGG 1385  
|||||  
Db 779 GAGAAGTCTGACGTTTCTCATTTGG 804  
|||||

RESULT 9

US-10-424-599-47702  
; Sequence 47702, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 47702  
; LENGTH: 1552  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_143080C.1  
US-10-424-599-47702

Query Match 1.3%; Score 26; DB 13; Length 1552;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1360 GAGAAGTCTGACGTTTCTCATTTGG 1385  
|||||  
Db 779 GAGAAGTCTGACGTTTCTCATTTGG 804  
|||||

RESULT 10  
US-10-425-114-2326  
; Sequence 2326, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 2326  
; LENGTH: 1724  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700209610\_FLI  
US-10-425-114-2326

Query Match 1.3%; Score 26; DB 13; Length 1724;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1231 TTGAAGCTAGGTTGCTGATTTGG 1256  
|||||  
Db 832 TTGAAGCTAGGTTGCTGATTTGG 857  
|||||

RESULT 11  
US-10-424-599-122411  
; Sequence 122411, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua

```

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 122411
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MBT3847_81542C.1
US-10-424-599-122411

```

```

Query Match      1.3%; Score 25; DB 13; Length 1101;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1178 CTAATAATCATTCACCGTGATATCAA 1202
      |||||||
DB 69 CTAATAATCATTCACCGTGATATCAA 93

```

```

RESULT 12
US-10-368-046-1
; Sequence 1, Application US/10368046
; Publication No. US20040063188A1
; GENERAL INFORMATION:
; APPLICANT: Robin A. Robinson
; APPLICANT: Victoria Cloce
; TITLE OF INVENTION: Method for Isolation and Purification of
; FILE OF INVENTION: Expressed Gene Products In Vitro
; FILE REFERENCE: 44149-3US1
; CURRENT APPLICATION NUMBER: US/10/368,046
; CURRENT FILING DATE: 2003-02-15
; PRIOR APPLICATION NUMBER: US 60/356,119
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,161
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,118
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,133
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,157
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,156
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,123
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,113
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,154
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,135
; PRIOR FILING DATE: 2002-02-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HPV-16 L1 codon optimized sequence
US-10-368-046-1

```

```

Query Match      1.2%; Score 24; DB 13; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 669 ACCTCTCCACCTCCACCACCGCC 692
      |||||||

```

```

DB 1471 ACCTCTCCACCTCCACCACCGCC 1494

RESULT 13
US-10-367-095-1
; Sequence 1, Application US/10367095
; Publication No. US20030228696A1
; GENERAL INFORMATION:
; APPLICANT: Robin A. Robinson
; TITLE OF INVENTION: No. US20030228696A1el Insect Cell Line
; FILE REFERENCE: 44149-1US1
; CURRENT APPLICATION NUMBER: US/10/367,095
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 60/356,119
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,161
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,118
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,133
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,157
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,156
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,123
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,113
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,154
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/356,135
; PRIOR FILING DATE: 2002-02-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HPV-16 L1 codon optimized sequence
US-10-367-095-1

Query Match      1.2%; Score 24; DB 16; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0.059;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 669 ACCTCTCCACCTCCACCACCGCC 692
      |||||||
DB 1471 ACCTCTCCACCTCCACCACCGCC 1494

RESULT 14
US-10-029-386-26494
; Sequence 26494, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26494
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

; OTHER INFORMATION: MAP TO CHR7.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.58  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.59  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45  
; OTHER INFORMATION: SWISSPROT HIT: P16056, EVALUE 1.00e-27  
; OTHER INFORMATION: NT HIT: G1455746, EVALUE 8.00e-90  
; OTHER INFORMATION: EST\_HUMAN HIT: AW418529.1, EVALUE 8.00e-97  
US-10-029-386-26494

Query Match 1.2%; Score 23; DB 15; Length 180;  
Best Local Similarity 100.0%; Pred. No. 0.2;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1240 AAGTTGCTGATTTGGTCTTGC 1262  
|||||  
Db 29 AAGTTGCTGATTTGGTCTTGC 51

RESULT 15  
US-10-029-386-12794  
; Sequence 12794, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 12794  
; LENGTH: 579  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR7.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.58  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.59  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.45  
; OTHER INFORMATION: NT HIT: U96996.1, EVALUE 1.00e-107  
; OTHER INFORMATION: SWISSPROT HIT: P16056, EVALUE 6.00e-27  
; OTHER INFORMATION: EST\_HUMAN HIT: AW418529.1, EVALUE 0.00e+00  
US-10-029-386-12794

Query Match 1.2%; Score 23; DB 15; Length 579;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1240 AAGTTGCTGATTTGGTCTTGC 1262  
|||||  
Db 221 AAGTTGCTGATTTGGTCTTGC 243

Search completed: May 12, 2004, 22:42:51  
Job time : 843 secs

**This Page Blank (uspto)**